```
                 :   |||:::          |||   |||||||||
Db       1165 pProLeuGlnTyrMetArgAlaAspGluAlaAlaGly---------------------- 1177

Qy        176 GCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCA-------- 125
                                                          |||       |||
Db       1178 ---------------------------------------GlyLeuArgProHisAspGl 1184

Qy        124 ---------GCTTCTCAGCCCATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTGGGG 75
                        |||:::          ||||||
Db       1184 uGluValGluAlaThrLeuLysSerLeuAsn--------------------------As 1195

Qy         74 ACACGTCTCATCACTCAGATCCTGGCCGAGGCGCGCGGCTGTCACCCGGAGCCAGCGCGT 15
                 :::   |||:::||||||      |||:::|||   :::|||   |||      |
Db       1195 nGlnIleGluSerIleArgSer-ProGluGlySerArgLysAsnPro---AlaArgThrC 1214

Qy         14 GCAGG 10
              |||||
Db       1214 ysArg 1215

RESULT 20
CGHU6C
collagen alpha 1(II) chain precursor [validated] - human
N;Alternate names: procollagen alpha 1(II) chain
N;Contains: chondrocalcin; collagen alpha 1(II) chain precursor splice form 1; collagen
C;Species: Homo sapiens (man)
C;Date: 28-May-1986 #sequence_revision 01-Sep-1995 #text_change 08-Dec-2000
C;Accession: A38513; S06715; S24270; A24828; S06496; A35428; A30147; A33116; S64674; S63
7250; I37251; I37252; I37253; I37254; I55338; I59535; I61910
R;Ryan, M.C.; Sieraski, M.; Sandell, L.J.
Genomics 8, 41-48, 1990
A;Title: The human type II procollagen gene: identification of an additional protein-cod
A;Reference number: A38513; MUID:91184811; PMID:2081599
A;Accession: A38513
A;Molecule type: DNA
A;Residues: 1-103 <RYA>
A;Cross-references: GB:M60299; NID:g180883; PIDN:AAA73873.1; PID:g180884
R;Su, M.W.; Lee, B.; Ramirez, F.; Machado, M.; Horton, W.
Nucleic Acids Res. 17, 9473, 1989
A;Title: Nucleotide sequence of the full length cDNA encoding for human type II procolla
A;Reference number: S06715; MUID:90067946; PMID:2587267
A;Accession: S06715
A;Molecule type: mRNA
A;Residues: 1-28,'R',99-1487 <SU2>
A;Cross-references: EMBL:X16468; NID:g29515; PIDN:CAA34488.1; PID:g29516
A;Note: alternative splice form 1
R;Vikkula, M.; Metsaeranta, M.; Syvaenen, A.C.; Ala-Kokko, L.; Vuorio, E.; Peltonen, L.
Biochem. J. 285, 287-294, 1992
A;Title: Structural analysis of the regulatory elements of the type-II procollagen gene.
A;Reference number: S24270; MUID:92344585; PMID:1637314
A;Accession: S24270
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-28 <VIK>
A;Cross-references: EMBL:X58709; GB:S40537; NID:g35659
A;Note: this translation is not annotated in GenBank entry HSPROCOE1, release 111.0
R;Nunez, A.M.; Kohno, K.; Martin, G.R.; Yamada, Y.
Gene 44, 11-16, 1986
A;Title: Promoter region of the human pro-alpha-1-(II)-collagen gene.
A;Reference number: A24828; MUID:87031574; PMID:3021582
A;Accession: A24828
A;Molecule type: DNA
A;Residues: 1-8,'T',10-28 <NUN>
A;Cross-references: GB:M25698; NID:g180872; PIDN:AAA52051.1; PID:g553237
R;Baldwin, C.T.; Reginato, A.M.; Smith, C.; Jimenez, S.A.; Prockop, D.J.
Biochem. J. 262, 521-528, 1989
A;Title: Structure of cDNA clones coding for human type II procollagen. The alpha-1(II)
A;Reference number: S06496; MUID:90026318; PMID:2803268
A;Accession: S06496
A;Molecule type: mRNA
A;Residues: 7-28,'R',99-157,'P',159-440,'G',442-456,'E',458-640,'A',642-831,'PA',834,'F'
A;Cross-references: EMBL:X16711; NID:g30040; PIDN:CAA34683.1; PID:g30041
A;Note: alternative splice form 1
R;Ryan, M.C.; Sandell, L.J.
J. Biol. Chem. 265, 10334-10339, 1990
A;Title: Differential expression of a cysteine-rich domain in the amino-terminal prop
A;Reference number: A35428; MUID:90285153; PMID:2355003
A;Accession: A35428
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 27-81,'L',83-103 <RYA2>
A;Note: alternative splice form 2; splicing appears to be under developmental regulat
R;Su, M.W.; Benson-Chanda, V.; Vissing, H.; Ramirez, F.
Genomics 4, 438-441, 1989
A;Title: Organization of the exons coding for Pro alpha-1(II) collagen N-propeptide c
A;Reference number: A30147; MUID:89233138; PMID:2714801
A;Accession: A30147
A;Molecule type: DNA
A;Residues: 104-157,'P',159-236 <SUM>
A;Cross-references: GB:J03065; GB:M23660; GB:M25655; GB:M25656; GB:M25730; GB:M32168;
R;Ala-Kokko, L.; Baldwin, C.T.; Moskowitz, R.W.; Prockop, D.J.
Proc. Natl. Acad. Sci. U.S.A. 87, 6565-6568, 1990
A;Title: Single base mutation in the type II procollagen gene (COL2A1) as a cause of
A;Reference number: A94227; MUID:90370826; PMID:1975693
A;Accession: A33116
A;Molecule type: DNA
A;Residues: 171-172,'C',174-175 <ALA>
A;Note: mutant sequence from a family with family with primary generalized osteoarthr
R;Diab, M.; Wu, J.J.; Eyre, D.R.
Biochem. J. 314, 327-332, 1996
A;Title: Collagen type IX from human cartilage: a structural profile of intermolecula
A;Reference number: S64673; MUID:96195147; PMID:8660302
A;Accession: S64674
A;Molecule type: protein
A;Residues: 188-189,'X',191-195;1224-1230,'X',1232-1236 <DIA>
R;Franc, S.; Marzin, E.; Boutillon, M.M.; Lafont, R.; Lechene de la Porte, P.; Herbag
Eur. J. Biochem. 234, 125-131, 1995
A;Title: Immunohistochemical and biochemical analyses of 20000-25000-year-old fossil
A;Reference number: S63514; MUID:96096730; PMID:8529631
A;Accession: S63514
A;Molecule type: protein
A;Residues: 243-261;575-590;756-763,'X',765-779 <FRA>
R;Tiller, G.E.; Weis, M.A.; Polumbo, P.A.; Gruber, H.E.; Rimoin, D.L.; Cohn, D.H.; Ey
Am. J. Hum. Genet. 56, 388-395, 1995
A;Title: An RNA-splicing mutation (G+5IVS20) in the type II collagen gene (COL2A1) in
A;Reference number: I38867; MUID:95150028; PMID:7847372
A;Accession: I38867
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 440,'G',442-456,'E',458-480,'P',482-509 <TIL1>
A;Cross-references: EMBL:U15195; NID:g557053; PIDN:AAB60370.1; PID:g557054
R;Ramirez, F.
submitted to the EMBL Data Library, December 1988
A;Reference number: S04892
A;Accession: S04892
A;Molecule type: mRNA
A;Residues: 501-676,'A',678-783,'A',785-831,'PA',834,'F',836-1214 <RAM>
A;Cross-references: EMBL:X13783; NID:g30037; PIDN:CAA32030.1; PID:g930050
R;Vikkula, M.; Peltonen, L.
FEBS Lett. 250, 171-174, 1989
A;Title: Structural analyses of the polymorphic area in type II collagen gene.
A;Reference number: S05000; MUID:89325561; PMID:2753125
A;Accession: S05000
A;Molecule type: DNA
A;Residues: 630-640,'A',642-785 <VIK2>
A;Cross-references: EMBL:X16158; NID:g29951; PIDN:CAA34278.1; PID:g1335018; PIDN:CAA3
 PIDN:CAA34283.1; PID:g1335023; PIDN:CAA34284.1; PID:g1335024
R;Bogaert, R.; Tiller, G.E.; Weis, M.A.; Gruber, H.E.; Rimoin, D.L.; Cohn, D.H.; Eyre
J. Biol. Chem. 267, 22522-22526, 1992
A;Title: An amino acid substitution (Gly853-->Glu) in the collagen alpha 1(II) chain
A;Reference number: A44309; MUID:93054548; PMID:1429602
A;Accession: A44309
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: DNA; mRNA
A;Residues: 752-831,'PA',834,'F',836-1005,'K',1007-1036,'Q',1038-1052,'E',1054-1068,'
A;Cross-references: GB:L00977; NID:g180812; PIDN:AAB23914.1; PID:g258774
A;Note: sequence extracted from NCBI backbone (NCBIP:117273); parts of this sequence
```

A;Note: this translation is not annotated and this publication is not cited in GenBank e
A;Note: mutant sequence associated with perinatal lethal hypochondrogenesis
R;Tiller, G.E.; Rimoin, D.L.; Murray, L.W.; Cohn, D.H.
Proc. Natl. Acad. Sci. U.S.A. 87, 3889-3893, 1990
A;Title: Tandem duplication within a type II collagen gene (COL2A1) exon in an individua
A;Reference number: S16502; MUID:90251662; PMID:2339128
A;Accession: S16502
A;Molecule type: DNA
A;Residues: 1164-1184,'GPSGKDGANGIPGPI',1185-1199 <TIL2>
A;Cross-references: EMBL:M37126; NID:g180808; PIDN:AAA52037.1; PID:g180809
A;Note: mutant sequence from a patient with spondyloepiphyseal dysplasia
R;Cheah, K.S.E.; Stoker, N.G.; Griffin, J.R.; Grosveld, F.G.; Solomon, E.
Proc. Natl. Acad. Sci. U.S.A. 82, 2555-2559, 1985
A;Title: Identification and characterization of the human type II collagen gene (COL2A1)
A;Reference number: A02858; MUID:85190534; PMID:3857598
A;Accession: A02858
A;Molecule type: DNA
A;Residues: 1032-1056,'N',1058-1068,'T',1070-1487 <CHE>
A;Cross-references: GB:J00116; NID:g180395; PIDN:AAA51997.1; PID:g180396
R;Elima, K.; Vuorio, T.; Vuorio, E.
Nucleic Acids Res. 15, 9499-9504, 1987
A;Title: Determination of the single polyadenylation site of the human pro-alpha-1(II) c
A;Reference number: A27280; MUID:88067771; PMID:2825137
A;Accession: A27280
A;Molecule type: DNA; mRNA
A;Residues: 1175-1487 <ELI>
A;Cross-references: EMBL:X06268; NID:g30096; PIDN:CAA29604.1; PID:g30097
A;Experimental source: fetal epiphyseal cartilage
R;van der Rest, M.; Rosenberg, L.C.; Olsen, B.R.; Poole, A.R.
Biochem. J. 237, 923-925, 1986
A;Title: Chondrocalcin is identical with the C-propeptide of type II procollagen.
A;Reference number: A57033; MUID:87099927; PMID:3800925
A;Accession: A57033
A;Molecule type: protein
A;Residues: 'XE',1244-1246,'N',1248,'X',1250-1265;1295-1305;1395-1408 <VAN>
A;Note: chondrocalcin identified as released collagen 1(II) chain carboxyl-terminal prop
R;Strom, C.M.; Upholt, W.B.
Nucleic Acids Res. 12, 1025-1038, 1984
A;Title: Isolation and characterization of genomic clones corresponding to the human typ
A;Reference number: A21733; MUID:84118798; PMID:6320112
A;Accession: A21733
A;Molecule type: DNA
A;Residues: 1245-1295 <STR1>
A;Cross-references: EMBL:X00339; EMBL:X00298; NID:g394699; PIDN:CAA25092.1; PID:g4378975
A;Accession: B21733
A;Molecule type: DNA
A;Residues: 894-909,'PE' <STR2>
A;Cross-references: GB:K01785; NID:g30035; PIDN:CAA25082.1; PID:g1335032
R;Nunez, A.M.; Francomano, C.; Young, M.F.; Martin, G.R.; Yamada, Y.
Biochemistry 24, 6343-6348, 1985
A;Title: Isolation and partial characterization of genomic clones coding for a human prc
gene.
A;Reference number: A24561; MUID:86104139; PMID:3002437
A;Accession: A24561
A;Molecule type: DNA
A;Residues: 1296-1358 <NUN2>
A;Cross-references: GB:M12048; NID:g180017
A;Note: this translation is not annotated in GenBank entry HUMCCT2A, release 111.0
A;Note: the codons given for 1333-Lys (AGG) and 1350-Gly (GCA) are inconsistent with the
R;Sangiorgi, F.O.; Benson-Chanda, V.; de Wet, W.J.; Sobel, M.E.; Tsipouras, P.; Ramirez,
Nucleic Acids Res. 13, 2207-2225, 1985
A;Title: Isolation and partial characterization of the entire human pro alpha 1(II) coll
A;Reference number: I37249; MUID:85215609; PMID:2987845
A;Accession: S59491
A;Molecule type: DNA
A;Residues: 7-28;'R',99-114;541-578;786-802;1055-1056,'N',1058-1068,'T',1070-1109;1200-1
A;Accession: I84453
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 7-28 <SAN2>
A;Cross-references: GB:M23759; NID:g180845; EMBL:X03320; GB:M24938; NID:g30104
A;Note: the GenBank PID is based on an incorrect reading frame
A;Accession: I37250
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 541-560 <SAN3>
A;Cross-references: EMBL:X02378; GB:M23870; NID:g30107; PIDN:CAA26227.1; PID:g929621
A;Accession: I37251

Alignment Scores:

Pred. No.:	2.89e-12	Length:	1487
Score:	328.00	Matches:	252
Percent Similarity:	31.02%	Conservative:	57
Best Local Similarity:	25.30%	Mismatches:	344
Query Match:	5.27%	Indels:	345
DB:	1	Gaps:	55

US-09-759-143-110 (1-3410) x CGHU6C (1-1487)

```
Qy     2588 TGTCTCTGTGATGGCAACAGAAGGACCAAC-----------AGGCCACATCCTGATAAA 2541
            |||:::  ||||||  |||       :::          :::|||  |||
Db       34 CysValGlnAspGlyGlnArgTyrAsnAspLysAspValTrpLysProGluPro------ 51

Qy     2540 AGGTAAGAGGGGGGTGGATCAGCAAAAAGACAGTGCTGTGGGCTGAGGGGACCTGGTTCT 2481
                                              |||  :::
Db       52 ----------------------------------CysArgIle-------------- 54

Qy     2480 TGTGTGTTGCCCCTCAGGACTCTTCCCCTACAAATAAGTCATATGTTCAAATCCCATGGA 2421
            |||||| |||   :::|||                 :::|||||||  :::   :::::
Db       55 CysVal-CysAspThrGlyThrValLeuCysAspAspIleIleCysGluAsp---ValLy 73

Qy     2420 GGAGTGTTTCATCCTAGAAACTCCCATGCAAGAGCTACATTAAACGAAGCTGCAGGTTAA 2361
            :::::|||        |||   |||    |||
Db       73 sAspCysLeuSerProGluIleProPheGlyGlu------------------------- 84

Qy     2360 GGGGCTTAGAGATGGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCT 2301
                                                 |||    |||:::
Db       85 -----------------------------------CysCysProIleCysProThr-- 91

Qy     2300 AGGTGTGTCTCAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCAGA 2241
                                        |||     :::   |||   |||
Db       92 ------------------------AspLeuAlaThrAlaSerGlyGlnPro-------- 100

Qy     2240 GTCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAAC 2181
                                 |||   |||
Db      101 ---------------------GlyProLysGly------------------------ 104

Qy     2180 CCCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCC 2121
                            |||  |||  |||  |||     :::     |||
Db      105 ----------------GlnLysGlyGluProGlyAspIleLysAspIleValGlyProLy 119

Qy     2120 CCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGAGCCACA 2061
              ||||||    |||                         |||   :::   |||
Db      119 sGlyProProGlyProGln---------------GlyProAlaGlyGluGlnGlyProAr 134

Qy     2060 TTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAACAGGAG 2001
                             ::::::   :::   |||     |||   |||     |||
Db      134 gGlyAspArgGlyAspLysGlyGluLysGlyAlaProGlyProArgGly------ArgAs 152

Qy     2000 CGGGGAGCTGGGA--------------------CCCAGTGAGGCAGGCCCT------CC 1968
             ||||||   |||                       |||       ||||||     ||
Db      152 pGlyGluProGlyThrLeuGlyAsnProGlyProProGlyProProGlyProProGlyPr 172

Qy     1967 ACCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATA 1908
             ||||        |||:::|||           :::   |||
Db      172 oProGlyLeuGlyGlyAsnPheAlaAlaGlnMetAlaGlyGly---------------- 186

Qy     1907 CTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGG 1848
                          |||      |||   |||      |||:::     |||    |||
Db      187 ----PheAspGluLysAlaGlyGlyAlaGlnLeu--GlyValMetGlnGlyPro------ 202

Qy     1847 CAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGG 1788
                               |||          |||            |||||||||||
Db      203 ------------------MetGlyProMetGlyPro---------ArgGlyProProGly 213

Qy     1787 ACAGCAGGAAGGCACTATCCA---GGATGGCGAGGT-----------CCAGGCAGATGC 1743
```

```
                 ||||||       |||  |||:::::|||            ||||||
Db        214 ProAlaGlyAlaProGlyProGlnGlyPheGlnGlyAsnProGlyGluProGlyGlu--- 232

Qy       1742 CCCGGCCCGGAACCACCCTGG---CCTCGGTGGGCTCACCCACCACCACA----CGTACG 1690
              ||||||         |||      ||||||          ||||||
Db        233 ProGlyValSerGlyProMetGlyProArgGlyProProGlyProProGlyLysProGly 252

Qy       1689 GAGACATCACAGGCAGAGGCCCCGCAGAGCGCGGGTGGAGGTGGAGCAGGCCACTGCCT 1630
              :::      :::|||       |||     ||||||     |||            |||
Db        253 AspAspGlyGluAlaGlyLysProGlyLysAlaGlyGluArgGly------------Pro 268

Qy       1629 CCAGCACCCACGTGTCCATTAGGG---AAGGGAGCTCCAGGCTTAGGGCCTGGCAGGAAG 1573
              |||   |||             |||    |||   |||||||||   ||||||   |||
Db        269 ProGlyProGlnGlyAlaArgGlyPheProGlyThrProGlyLeu---ProGlyValLys 287

Qy       1572 CTGGTCATCAGGCTGTCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGG 1513
                ||||||:::
Db        288 --GlyHisArgGly------------------------------------------- 291

Qy       1512 AACACCTGCTTCTCCCGGTGGTAGAGGGA------GGCCAGTGTGTAGGGCAGGATCTGC 1459
                                |||   ::::::|||    |||:::            ::::
Db        292 ----------TyrProGlyLeuAspGlyAlaLysGlyGluAlaGlyAlaProGlyValL 308

Qy       1458 AGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGG 1399
              ::|||      |||    ||||||||    |||                    :::   |
Db        308 ysGlyGluSerGlySerProGlyGluAsnGly--------------------SerProG 321

Qy       1398 GACAGGCATGTGGCACCGGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCT 1339
              ||                 |||         |||       ||||||              :
Db        321 lyProMetGlyProArgGlyLeuProGlyGluArgGlyArgThrGlyProAlaGlyAlaA 341

Qy       1338 CGAGTGCCGAATCGCT-------------GCACCAGCCGGTCCATGACCAGAGAGAAGAC 1292
              ::   |||                        ||||||||||||               |
Db        341 laGlyAlaArgGlyAsnAspGlyGlnProGlyProAlaGlyPro--------------P 356

Qy       1291 CAGGGAGATGGCGCACTGCAGGAACAGCC------CCAGGCTGCCCATCCGAACGCCTTC 1238
              |||||             ||||||         ||||||   |||         :::
Db        356 roGlyProValGlyProAlaGlyGlyProGlyPheProGlyAlaProGlyAlaLysGlyG 376

Qy       1237 ATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTACAGCCC 1178
                        :::|||   |||   |||                     |||
Db        376 luAlaGlyProThrGlyAlaArgGlyProGluGlyAlaGlnGlyProArgGlyGluProG 396

Qy       1177 CTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGCTGCACAG 1118
                |||   :::|||                        |||:::         :::
Db        396 lyThrProGlySerPro-----------------------GlyProAlaGlyAlaSerG 408

Qy       1117 CTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACAGCTGGTGCAGCCG 1058
                   |||        |||        ||||||   ||||||   |||||||||   |
Db        408 lyAsnProGlyThrAspGlyIlePro---GlyAlaLysGlySer---AlaGlyAlaProG 426

Qy       1057 GGGAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGGCCCGGCATGGACAGCAGTGGGG 998
              ||   ||||||   ||||||        |||   ||||||      |||        ||||
Db        426 lyIleAlaGlyAlaProGlyPheProGlyProArgGlyPro---ProAspProGlnGlyA 445

Qy        997 CGACA---------------------------------------------------- 993
              |||||
Db        445 laThrGlyProLeuGlyProLysGlyGlnThrGlyLysProGlyIleAlaGlyPheLysG 465

Qy        992 --AGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCTCCT---- 939
                 :::||||||                          |||||||||   |||
Db        465 lyGluGlnGlyProLys-----------GlyGluProGlyProAlaGlyProGlnGlyA 481

Qy        938 -----------------------------------------------CAGCCACCAGCAGTGT 923
                                                             :::|||   |||
Db        481 laProGlyProAlaGlyGluGluGlyLysArgGlyAlaArgGlyGluProGlyGlyValG 501

Qy        922 GGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGC------------- 876
                    :::         |||    |||    |||    ||||||
Db        501 lyProIleGlyProProGlyGluArgGlyAlaProGlyAsnArgGlyPheProGlyGlnA 521

Qy        875 ----------ACTCCTCCTGGGTGCCCAGGT---AGGGGGCCAGGGCACTGGTGTCCCA 830
                        |||   |||   ||||||    |||||||||
Db        521 spGlyLeuAlaGlyProLysGlyAlaProGlyGluArgGlyProSerGlyLeuAlaGlyP 541

Qy        829 GTCAATGGCAGGCAGGA------GGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGC 776
               :::            |||     |||    ||||||   |||               |||
Db        541 roLysGlyAlaAsnGlyAspProGlyArgProGlyGluProGlyLeuProGlyAlaArgG 561

Qy        775 ATAGACAGAGTAGGCCTGGCGACAGTGGTCCGGGTCCCGGAAGAGGTCAGAGAGCAGGGC 716
                          ||||||||||||:::||||||     |||:::          :::|||
Db        561 lyLeuThrGlyArgProGlyAspAlaGlyPro---GlnGlyLysValGlyProSerGlyA 580

Qy        715 CTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGC---------- 666
                |||   |||          ||||||         |||         |||
Db        580 laProGlyGluAspGlyArgProGlyProProGlyProGlnGlyAlaArgGlyGlnProG 600

Qy        665 -------------CCAGGATGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGCACAG 620
                          ||||||          |||                  :::||||||   |
Db        600 lyValMetGlyPheProGly---------ProLysGlyAlaAsnGlyGluProGlyLysA 617

Qy        619 CAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCA---- 564
              ||                                   |||:::|||      |||   |||
Db        617 laGly-------------------------GluLysGlyLeuProGlyAlaProGlyL 628

Qy        563 --AGGACAGTGCCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACT 506
                |||   :::|||                                          |||
Db        628 euArgGlyLeuPro----------------------------------------GlyLysA 635

Qy        505 GGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCA---------------------C 467
                       ||||||            ||||||||||||
Db        635 spGlyGluThrGlyAlaGluGlyProProGlyProAlaGlyProAlaGlyGluArgGlyG 655

Qy        466 TGGACCAATGCCCAGCACCATGGTCATGAACTTCTCCTCTACCCCCACTTCCAGCAGCAG 407
                   |||   |||                   |||        ||||||
Db        655 luGlnGlyAlaProGlyPro------------SerGlyPheGlnGlyLeuProGly---- 669

Qy        406 AGGCGGCACATAGGTGATGCCTGCGGCCAAACACACCTCCAGGCCAAAGGTTAGCAGGTT 347
                                        |||     |||||||||:::        |||
Db        670 -------------------------ProProGlyProProGlyGlu------GlyGlyL 679

Qy        346 GACCAGCAAGAGCTGGGCTTTCCGGTGCCGCAGCAGGC--------------------- 309
                |||        |||:::   |||   |||
Db        679 ysProGlyAspGlnGlyValProGlyGluAlaGlyAlaProGlyLeuValGlyProArgG 699

Qy        308 -------------------------GGCTCACCCACAGCCTCTGGA------------- 288
                                       |||||||||   |||   |||
Db        699 lyGluArgGlyPheProGlyGluArgGlySerProGlyAlaGlnGlyLeuGlnGlyProA 719

Qy        287 --------CCATAGTGGGCCAGGCGGGTAGGGCTCAGGGGGCCGTTCAGGCACTCCAGAA 236
                         |||     ||||||         |||       |||    ||||||
Db        719 rgGlyLeuPro-GlyThrProGlyThrAspGlyProLysGlyAlaSerGlyPro------ 736

Qy        235 CTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAACTG 176
                                   :::    |||      |||         ||||||      :::   :::
Db        737 ---------------AlaGlyProProGlyAlaGlnGlyProProGlyLeuGlnGlyMet 751

Qy        175 CCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGT----------------- 134
              |||       :::         |||                 |||   |||
Db        752 ProGlyGluArgGlyAlaAlaGlyIleAlaGlyProLysGlyAspArgGlyAspValGly 771

Qy        133 ---GCCGGTCCAGCTTCTCAGCCC-----------------------------ATGCTC 107
                 ||||||         |||                                  :::
Db        772 GluLysGlyProGluGlyAlaProGlyLysAspGlyGlyArgGlyLeuThrGlyProIle 791

Qy        106 AACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTCATCACTCAGATCCTGGC-- 49
                    |||                        :::|||:::               ||||||
Db        792 Gly-ProProGlyProAlaGlyAlaAsnGlyGluLysGlyGluValGlyProProGlyPr 811

Qy         48 ---------CGAGGCGCGCGGCTGTCACCCGGA 25
                       |||||||||    :::||||||
Db        811 oAlaGlySerAlaGlyAlaArgGlyAlaProGly 822
```

RESULT 21
S28052
sucrose transport protein - spinach

```
N;Alternate names: sucrose carrier protein; sucrose permease
C;Species: Spinacia oleracea (spinach)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 21-Jul-2000
C;Accession: S28052
R;Riesmeier, J.W.; Willmitzer, L.; Frommer, W.B.
EMBO J. 11, 4705-4713, 1992
A;Title: Isolation and characterization of a sucrose carrier cDNA from spinach by functi
A;Reference number: S28052; MUID:93099843; PMID:1464305
A;Accession: S28052
A;Molecule type: mRNA
A;Residues: 1-525 <RIE>
A;Cross-references: EMBL:X67125; NID:g21318; PIDN:CAA47604.1; PID:g21319
C;Superfamily: common tobacco sucrose transport protein
C;Keywords: transmembrane protein

Alignment Scores:
Pred. No.:              3.32e-12     Length:         525
Score:                  327.00       Matches:        134
Percent Similarity:     39.58%       Conservative:   92
Best Local Similarity:  23.47%       Mismatches:     200
Query Match:            5.10%        Indels:         145
DB:                     2            Gaps:           21

US-09-759-143-110 (1-3410) x S28052 (1-525)

Qy        310 CCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGA 369
              |||          |||||  |||:::        |||      |||          ||||||
Db         25 ProThrThrProProGlu-AlaGluAlaThrLeuLysLysLeu---------GlyLeuVa 41

Qy        370 GGTGTGTTTGGCCGCAGGCATCACCTAT-----------------------GTGCCGCC 405
                     :::||||||||:::   :::                          :::   ||
Db         41 lAlaSerValAlaAlaGlyValGlnPheGlyTrpAlaLeuGlnLeuSerLeuLeuThrPr 61

Qy        406 TCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCC 465
               |   :::        :::|||:::          :::        :::        |||||
Db         61 oTyrValGlnLeuLeuGlyIleProHisThrTrpAlaAlaTyrIleTrpLeuCysGlyPr 81

Qy        466 AGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACG 525
              |:::   |||:::::::        ||||||:::|||       ||||||          ||
Db         81 oIleSerGlyMetIleValGlnProLeuValGlyTyrTyrSerAspArgCysThrSerAr 101

Qy        526 CTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTT 585
              |:::||||||||||||||||||||||||   |||    :::      :::    :::::::::
Db        101 gPheGlyArgArgArgProPheIleAlaAlaGlyAlaAlaLeuValAlaValAlaValGl 121

Qy        586 TCTCATCCCAAGGGCCGGCTGGCTAGCA------GGGCTGCTGTGCCCGGATCCC----- 634
                ||||||         |||:::   |||      |||              ||||||
Db        121 yLeuIle---------GlyPheAlaAlaAspIleGlyAlaAlaSerGlyAspProThrGl 138

Qy        635 ----------AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTT 684
                          :::|||    :::|||:::     :::::|||      :::||||||
Db        138 yAsnValAlaLysProArgAlaIleAlaValPheValValGlyPheTrpIleLeuAspVa 158

Qy        685 CTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGG---GA 741
                                      |||      ||||||||:::|||:::
Db        158 lAlaAsnAsnThrLeuGlnGlyProCysArgAlaLeuLeuAlaAspMetAlaAlaGlySe 178

Qy        742 CCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTG 801
                            |||   |||       :::     ::::::|||         ::::::||||||
Db        178 rGlnThrLysThrArgTyrAlaAsnAlaPhePheSerPhePheMetAlaLeuGlyAsnIl 198

Qy        802 CCTGGGCTACCTCCTGCCTGCCATTGAC-----------------------TGGGACAC 837
                       ||||||            :::                                  ||
Db        198 eGlyGlyTyrAlaAlaGlySerTyrSerArgLeuTyrThrValPheProPheThrLysTh 218

Qy        838 CAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACC-- 895
              |:::|||             |||               :::    |||   |||    :::|||
Db        218 rAlaAlaCysAspValTyrCysAlaAsnLeuLysSerCysPhePheIleSerIleThrLe 238

Qy        896 -CTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGG 954
               ||||||   ||||||   :::|||   :::::::::           :::       :::
Db        238 uLeuIleValLeuThrIleLeuAlaLeuSerValValLysGluArgGlnIleThrIleAs 258

Qy        955 CCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATG 1014
                         :::       |||   |||            :::    |||       |||
Db        258 pGluIleGlnGluGluGluAspLeuLysAsnArgAsnAsnSerSerGlyCys-------- 275

Qy       1015 CCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTG 1074
                   |||||||||    |||              |||   |||:::          |||    |||
Db        276 ----AlaArgLeuProPhe------PheGlyGlnLeuIleGlyAlaLeuLysAspLeu-- 291

Qy       1075 CTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGC 1134
                              |||:::    :::       |||    :::          :::|||:::||
Db        292 ----------ProLysProMetLeuIleLeuLeuLeuValThrAlaLeuAsnTrpIleAl 308

Qy       1135 ACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGT 1194
              |          |||    ||||||    ||||||:::::|||:::      :::|||    |||
Db        308 aTrpPheProPheLeuLeuPheAspThrAspTrpMetGlyLysGluValTyrGlyGlyTh 328

Qy       1195 GCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGG 1254
                                              :::    ||||||:::||||||          ||
Db        328 r------------------ValGlyGluGlyLysLeuTyrAspGlnGlyValHisAlaGl 342

Qy       1255 CAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCG 1314
              |:::||||||||||    :::          :::         |||    |||||    ::::::
Db        342 yAlaLeuGlyLeuMetIleAsnSerValValLeuGlyValMetSerLeuSerIleGluGl 362

Qy       1315 GCTGGTGCAGCGATTCGGC------ACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 1368
               |||    :::       |||              |||           :::    ::::::
Db        362 yLeuAlaArgMetValGlyGlyAlaLysArgLeuTrpGlyIleValAsnIleIleLeu-- 381

Qy       1369 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 1428
                                  |||   ||||||:::    :::|||    :::||||||    |||||
Db        382 ----------------AlaValCysLeuAlaMetThrVal---LeuValThrLysSerAl 395

Qy       1429 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 1488
              |
Db        395 aGlu-------------------------------------------------------- 396

Qy       1489 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 1548
               :::   |||:::
Db        397 -HisPheArgAspSerHisHisIleMet-------------------------------- 405

Qy       1549 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 1608

Db        405 ------------------------------------------------------------ 405

Qy       1609 TAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 1668
                                    ||||||    :::    ||||||||||||||          ||
Db        406 ---------------------GlySerAlaValProProProProProAla------Gl 416

Qy       1669 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 1728
              |         :::::::::      |||:::|||    |||
Db        416 yValLysGlyGlyAlaLeuAlaIlePheAlaValLeuGlyIleProLeuAlaIleThrPh 436

Qy       1729 G---GTTCCG---------------------------------GGCCGGGGCATCTGCCT 1752
                 :::|||                                   |||    |||       ||
Db        436 eSerIleProPheAlaLeuAlaSerIlePheSerAlaSerSerGlySerGlyGlnGlyLe 456

Qy       1753 GGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTG----------------- 1795
              |    |||    :::|||:::    |||    ::::::    |||:::
Db        456 uSerLeuGlyValLeuAsnLeuAlaIleValValProGlnMetPheValSerValThrSe 476

Qy       1796 --------------------------------GCCCCATCCCTGTTTATGGGCTCCAT 1821
                                                    |||:::       :::|||:::::
Db        476 rGlyProTrpAspAlaMetPheGlyGlyGlyAsnLeuProAlaPheValValGlyAlaVa 496

Qy       1822 TGTCCAGCTCAGCCAGTCTGTCACTGCCTAT 1852
              :         :::    :::|||    ::::::
Db        496 lAlaAlaThrAlaSerAlaValLeuSerPhe 506

RESULT 22
S43142
sucrose transport protein - castor bean
N;Alternate names: sucrose carrier
C;Species: Ricinus communis (castor bean)
```

```
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 26-May-2000
C;Accession: S43142
R;Weig, A.; Komor, E.
submitted to the EMBL Data Library, March 1994
A;Description: A sucrose carrier from Ricinus communis.
A;Reference number: S43142
A;Accession: S43142
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-533 <WEI>
A;Cross-references: EMBL:Z31561; NID:g468561; PID:g468562
C;Superfamily: common tobacco sucrose transport protein

Alignment Scores:
Pred. No.:                3.81e-12      Length:        533
Score:                    326.00        Matches:       147
Percent Similarity:       40.75%        Conservative:  102
Best Local Similarity:    24.06%        Mismatches:    236
Query Match:              5.08%         Indels:        126
DB:                       2             Gaps:          22

US-09-759-143-110 (1-3410) x S43142 (1-533)

Qy      231 AGCAGTTCTGGAGTGCCTGAACGGCCCCCT------GAGCCCTACCCGCCTGGCCCACTA 284
            |||:::|||              ::::::||||||      :::|||:::||||||    ||||||
Db        4 SerThrSerLysGluAsnLysGlnProProSerSerGlnProHisProPro---ProLeu 22

Qy      285 TGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGG 344
                          |||       |||   |||    ::::::|||     ::::::::::::|
Db       23 MetValAlaGly-AlaAlaGlu---ProAsnSerSerPro-LeuArgLysValValMetV 41

Qy      345 TCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGC 404
            ||      :::       |||::::::        |||       ::::::    :::    |
Db       41 alAlaSerIleAlaAlaGlyIleGlnPheGlyTrpAlaLeuGlnLeuSerLeuLeuThrP 61

Qy      405 CTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTC 464
            ||     :::      :::|||:::        :::          :::        ||||
Db       61 roTyrValGlnLeuLeuGlyIleProHisThrTrpAlaAlaPheIleTrpLeuCysGlyP 81

Qy      465 CAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGAC 524
            ||:::    |||::::::         |||::::::|||       ||||||        |
Db       81 roIleSerGlyMetLeuValGlnProIleValGlyTyrHisSerAspArgCysThrSerA 101

Qy      525 GCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCT 584
            ||::|||||||||||||||||||||||      :::   :::      :::::::::
Db      101 rgPheGlyArgArgArgProPheIleAlaSerGlyAlaAlaPheValAlaIleAlaValP 121

Qy      585 TTCTCATCCCAAGGGCC---------GGCTGGCTAGCAGGG---CTGCTGTGCCCGGATC 632
            ||||||||      |||         |||   |||:::|||      |||           |
Db      121 heLeuIleGlyTyrAlaAlaAspLeuGlyHisLeuSerGlyAspSerLeuAspLysSerP 141

Qy      633 CCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCC 692
            ||:::       :::|||:::    ::::::|||      :::||||||
Db      141 roLysThrArgAlaIleAlaIlePheValValGlyPheTrpIleLeuAspValAlaAsnA 161

Qy      693 AGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTC---CGGGACCCGGACC 749
              :::            |||        |||||||||::::||||||
Db      161 snMetLeuGlnGlyProCysArgAlaLeuLeuAlaAspLeuSerGlyThrSerGlnLysL 181

Qy      750 ACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCT 809
                  |||   |||   :::::::::::::|||  :::::::::|||        |||||||
Db      181 ysThrArgThrAlaAsnAlaLeuPheSerPhePheMetAlaValGlyAsnValLeuGlyT 201

Qy      810 ACCTCCTGCCTGCCATTGAC-----------------------TGGGACACCAGTGCCC 845
            ||          |||                                   |||:::|||
Db      201 yrAlaAlaGlyAlaTyrThrHisLeuTyrLysLeuPheProPheThrLysThrThrAlaC 221

Qy      846 TGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCC 905
                     |||              :::   |||   |||       :::   |||:::
Db      221 ysAspValTyrCysAlaAsnLeuLysSerCysPhePheIleSerIleValLeuLeuLeuS 241

Qy      906 TCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGC 965
                     |||    |||          |||   |||:::            |||       |
Db      241 erLeuThrValLeuAlaLeuSerTyrValLysGluLys-----------ProTrpSerP 257

Qy      966 CAGCAGAAGGGCTG----------------------TCGGCCCCCTCCTTGTCGCCCC 1001
            ||   :::   :::                        |||       |||    |||
Db      257 roAspGlnAlaValAspAsnAlaGluAspAspThrAlaSerGlnAlaSerSerSerAlaG 277

Qy     1002 ACTGCTGTCCATGCCGGGCCCGCTTG------GCTTTCCGGAACCTGGGCGCCCTGCTTC 1055
                  |||             :::      ||||||:::||||||
Db      277 lnProMetProPhePheGlyGluIleLeuGlyAlaPheLysAsnLeu------------ 292

Qy     1056 CCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTG 1115
                                          |||   :::      |||   :::
Db      293 ---------------------------LysArgProMetTrpIleLeuLeuValT 303

Qy     1116 AGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTACACGGATTTCGTGGGCG 1175
                  :::|||:::|||              |||   ||||||   ||||||:::::|||
Db      303 hrCysLeuAsnTrpIleAlaTrpPheProPheLeuLeuPheAspThrAspTrpMetGlyA 323

Qy     1176 AGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATG 1235
                 :::|||   |||          :::             |||      :::   ||||
Db      323 rgGluValTyrGlyGly------AspSerSerGlySerAlaGluGlnLeuLysLeuTyrA 341

Qy     1236 ATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCT 1295
            ||   |||||||||   |||:::|||||||||   |||            :::
Db      341 spArgGlyValArgAlaGlyAlaLeuGlyLeuMetLeuAsnSerValValLeuGlyPheT 361

Qy     1296 TCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCA 1355
                ||||||   ::::::   |||   :::         |||               |||
Db      361 hrSerLeuGlyValGluValLeuAlaArgGlyValGly---GlyValLysArgLeuTrpG 380

Qy     1356 GTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGG 1415
              :::       |||    :::             |||   ||||||:::    :::|||    :::|
Db      380 lyIleValAsnPheValLeu---------AlaValCysLeuAlaMetThrVal---LeuV 396

Qy     1416 TGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACA 1475
              |||||      |||    |||
Db      396 alThrLysGlnAlaGluSerThr----------------------------------- 403

Qy     1476 CACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACA 1535
                                                         ::::::          :
Db      404 -----------------------------------------ArgArgPheAlaThrValS 410

Qy     1536 CTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTG 1595
             ::|||||||||                        ::::|||
Db      410 erGlyGlyAlaLys-------------------------ValPro-------------- 416

Qy     1596 GAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCAC 1655
                                                           |||||||||||
Db      417 --------------------------------------------LeuProProPro- 420

Qy     1656 CCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCA 1715
                        |||     |||     :::::::::        |||:::|||    |||
Db      421 --------SerGlyValLysAlaGlyAlaLeuAlaLeuPheAlaValMetGlyValProG 438

Qy     1716 CCGAGGCCAGGGTG---GTTCCG-------------------------------GGCC 1739
                           :::|||                                   |||
Db      438 lnAlaIleThrTyrSerIleProPheAlaLeuAlaSerIlePheSerAsnThrSerGlyA 458

Qy     1740 GGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCC 1799
              |||      |||    |||   :::|||:::   :::    ::::::   |||:::
Db      458 laGlyGlnGlyLeuSerLeuGlyValLeuAsnLeuSerIleValIleProGlnMetIleV 478

Qy     1800 CATCCCTGTTTATGGGCTCCATTGTCCAGCTC------AGCCAGTCTGTCACTGCCTATA 1853
              |||:::        |||          |||                ::::::   |||:::
Db      478 alSerValAlaAlaGlyProTrpAspAlaLeuPheGlyGlyGlyAsnLeuProAlaPheV 498

Qy     1854 TGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTG 1913
             ::|||    |||                          |||:::    |||    :::
Db      498 alValGlyAlaValAlaAlaLeuAlaSerGlyIlePheAlaLeuThrMetLeuProSerP 518

Qy     1914 ACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              :::    |||:::            |||
Db      518 roGlnProAspMetProSerAlaLysAla 527
```

RESULT 23
CGHU2V
collagen alpha 2(V) chain precursor - human
C;Species: Homo sapiens (man)
C;Date: 31-Jul-1989 #sequence_revision 28-Jul-1995 #text_change 31-Dec-2000
C;Accession: A31427; A54555; S43643; A25874; I55239; I59025; A25374; A30017
R;Woodbury, D.; Benson-Chanda, V.; Ramirez, F.
J. Biol. Chem. 264, 2735-2738, 1989
A;Title: Amino-terminal propeptide of human pro-alpha2(V) collagen conforms to the struc
A;Reference number: A31427; MUID:89123368; PMID:2914927
A;Accession: A31427
A;Molecule type: mRNA
A;Residues: 1-463 <WOO>
A;Cross-references: GB:J04478; NID:g179697; PIDN:AAA51859.1; PID:g179698
A;Experimental source: placenta
R;Greenspan, D.S.; Lee, S.T.; Lee, B.S.; Hoffman, G.G.
Gene Expr. 1, 29-39, 1991
A;Title: Homology between alpha2(V) and alpha1(III) collagen promoters and evidence for
A;Reference number: A54555; MUID:92314691; PMID:1820205
A;Accession: A54555
A;Molecule type: DNA
A;Residues: 1-32 <GRE>
A;Cross-references: GB:M58529; NID:g180834; PIDN:AAC41699.1; PID:g553235
R;Moradi-Ameli, M.; Rousseau, J.C.; Kleman, J.P.; Champliaud, M.F.; Boutillon, M.M.; Ber
Eur. J. Biochem. 221, 987-995, 1994
A;Title: Diversity in the processing events at the N-terminus of type-V collagen.
A;Reference number: S43642; MUID:94237164; PMID:8181482
A;Accession: S43643
A;Molecule type: protein
A;Residues: 288-291,'P',293-294,'X',296-297;606,'X',608-617 <MOR>
R;Weil, D.; Bernard, M.; Gargano, S.; Ramirez, F.
Nucleic Acids Res. 15, 181-198, 1987
A;Title: The pro alpha 2(V) collagen gene is evolutionarily related to the major fibril
A;Reference number: A25874; MUID:87146331; PMID:3029669
A;Accession: A25874
A;Molecule type: mRNA; DNA
A;Residues: 398-1496 <WEI>
A;Cross-references: GB:X04758; NID:g29588; PIDN:CAA28454.1; PID:g1340175
A;Experimental source: rhabdomyosarcoma cell line
R;Myers, J.C.; Loidl, H.R.; Stolle, C.A.; Seyer, J.M.
J. Biol. Chem. 260, 5533-5541, 1985
A;Title: Partial covalent structure of the human alpha 2 type V collagen chain.
A;Reference number: I55239; MUID:85182703; PMID:2985598
A;Accession: I55239
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1002-1226 <RE2>
A;Cross-references: GB:M10956; NID:g180427; PIDN:AAA52007.1; PID:g180428
A;Note: part of this sequence were determined by protein sequencing
R;Emanuel, B.S.; Cannizzaro, L.A.; Seyer, J.M.; Myers, J.C.
Proc. Natl. Acad. Sci. U.S.A. 82, 3385-3389, 1985
A;Title: Human alpha 1(III) and alpha 2(V) procollagen genes are located on the long arm
A;Reference number: I59025; MUID:85216505; PMID:3858826
A;Accession: I59025
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1003-1034 <RES>
A;Cross-references: GB:M11135; NID:g179693; PIDN:AAA51857.1; PID:g179694
A;Note: part of this sequence were determined by protein sequencing
R;Myers, J.C.; Loidl, H.R.; Seyer, J.M.; Dion, A.S.
J. Biol. Chem. 260, 11216-11222, 1985
A;Title: Complete primary structure of the human alpha-2 type V procollagen COOH-termina
A;Reference number: A25374; MUID:85289337; PMID:2411731
A;Accession: A25374
A;Molecule type: mRNA
A;Residues: 1227-1417,'T',1419-1437,'S',1439-1496 <MYE>
A;Cross-references: GB:M11718; NID:g180912; PIDN:AAA52058.1; PID:g180913
A;Experimental source: normal fibroblasts
R;Tsipouras, P.; Schwartz, R.C.; Liddell, A.C.; Salkeld, C.S.; Weil, D.; Ramirez, F.
Genomics 3, 275-277, 1988
A;Title: Genetic distance of two fibrillar collagen loci, COL3A1 and COL5A2, located on
A;Reference number: A30017; MUID:89138450; PMID:3224983
A;Accession: A30017
A;Molecule type: DNA
A;Residues: 1449-1463,'E',1465-1495,'A' <TSI>
A;Cross-references: GB:J03051; NID:g179695; PIDN:AAA51858.1; PID:g179696
A;Note: the authors translated the codon GAA for residue 1460 as Gln, and GAG for res
C;Comment: Prolines and lysines at the third position of the tripeptide repeating uni
are 5-hydroxylated and subsequently O-glycosylated.
C;Comment: The amino-terminal propeptide domain appears not to be completely cleaved.
C;Genetics:
A;Gene: GDB:COL5A2
A;Cross-references: GDB:119064; OMIM:120190
A;Map position: 2q31-2q31
A;Introns: 33/1; 812/3; 830/3; 848/3; 884/3; 902/3; 922/3; 974/3; 1046/3; 1064/3; 144
C;Complex: type V collagen may be a homotrimer of alpha 1(V) chains (see PIR:CGHU1V),
alpha 2(V) chain and one alpha 3(V) chain, initially linked by disulfide bonds among
ength, is formed with desmosine cross-links made from lysine and allysine residues
C;Function:
A;Description: structural component of extracellular fibrous polymer associated with
A;Note: may play a role in controlling the lateral growth of collagen I fibrils
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo
C;Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxypr
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-1250/Product: collagen alpha 2(V) chain #status predicted <MAT>
F;27-193/Domain: amino-terminal propeptide (uncleaved) #status predicted <NPP>
F;27-108/Region: nonhelical
F;40-99/Domain: von Willebrand factor type C repeat homology <VWC>
F;109-186/Region: helical
F;187-208/Region: nonhelical
F;209-1225/Region: helical
F;503-505/Region: cell attachment (R-G-D) motif
F;941-943/Region: cell attachment (R-G-D) motif
F;1064-1066/Region: cell attachment (R-G-D) motif
F;1067-1069/Region: cell attachment (R-G-D) motif
F;1097-1099/Region: cell attachment (R-G-D) motif
F;1124-1126/Region: cell attachment (R-G-D) motif
F;1133-1135/Region: cell attachment (R-G-D) motif
F;1225-1250/Region: carboxyl-terminal nonhelical telopeptide
F;1251-1496/Domain: carboxyl-terminal propeptide #status predicted <CPP>
F;1269-1496/Domain: fibrillar collagen carboxyl-terminal homology <FCC>
F;27/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic
F;193-194/Cleavage site: Ala-Gln (procollagen N-endopeptidase) #status predicted
F;194/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predi
F;201/Modified site: allysine (Lys) #status predicted
F;290,293,296,608,614,1004,1007,1013,1028,1034/Modified site: 4-hydroxyproline (Pro)
F;299,1139/Modified site: 5-hydroxylysine (Lys) #status predicted
F;299,1139/Binding site: carbohydrate (Lys) (covalent) #status predicted
F;1025/Modified site: 5-hydroxylysine (Lys) #status experimental
F;1250-1251/Cleavage site: Glu-Asp (procollagen C-endopeptidase) #status predicted
F;1259,1397/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;1293,1299,1325/Disulfide bonds: interchain #status predicted
F;1333-1494,1402-1447/Disulfide bonds: #status predicted

```
Alignment Scores:
Pred. No.:              4.7e-12     Length:        1496
Score:                  324.50      Matches:       304
Percent Similarity:     29.16%      Conservative:  45
Best Local Similarity:  25.40%      Mismatches:    390
Query Match:            5.21%       Indels:        458
DB:                     1           Gaps:          68

US-09-759-143-110 (1-3410) x CGHU2V (1-1496)

Qy       3068 GGTCCTGCAGTAGCTCCAAACAGGGTTGTGGAGCTGGTGGGGAAAGTTGGGGGTAGGGGA 3009
                 ||||||   |||   |||                 |||   |||:::   ||||||   |||
Db        216 GlyPro---ValGlyProArgGlyProGlnGlyLeuGlnGlyGlnGlnGlyGlyAlaGly 234

Qy       3008 AAGTTGGGG---GTAGGGGAAATTTTGGGCAGTGCCTTCATCAGCCCAGTCCTAGAGAGA 2952
                    |||      ||||||                      :::   |||:::      |||
Db        235 ProThrGlyProProGlyGluProGlyAspProGlyProMetGlyProIleGlySerArg 254

Qy       2951 ---GTAGAGGGG--------------------AGTGGAAGTGGGGGAACCAGGCTGGG 2916
                    ||||||                      |||      |||
Db        255 GlyProGluGlyProProGlyLysProGlyGluAspGlyGluProGlyArgAsnGlyAsn 274
```

```
Qy   2915 CCAAGAGAAGAGGGGTGGTTAGGGAAGCCGTTGAGA---------CCTGAAGCCCCACCC 2865
Db    275 ProGlyGluValGlyPheAlaGlySerProGlyAlaArgGlyPheProGlyAlaPro-Gl 294

Qy   2864 TCTACCTTCCTTCAACACCCTAACCTTGGGTAACAGCATTTGGAATTATCATTTGGGATG 2805
Db    294 yLeuPro--------------------GlyLeuLysGlyHisArgGlyHisLysGly-- 306

Qy   2804 AGTAGAATTTCCAAGGTCCTGGGTTAGGCATTTTGGG---------------------- 2768
Db    307 -LeuGluGlyProLysGlyGluValGlyAlaProGlySerLysGlyGluAlaGlyProTh 326

Qy   2767 ------------------GGGCCA---GACCCCAGGAGAAGAAGATTCTGGCAATGATC 2730
Db    326 rGlyProMetGlyAlaMetGlyProLeuGlyProArgGlyMet----------------- 340

Qy   2729 AGCCCAATGACCAGCTATCTCAGGGGACCTGATTGTTGGGGATCCCCCACCCTACCCAAA 2670
Db    341 ---------ProGlyGluArgGlyArgLeu---------GlyProGlnGlyAlaProGl 354

Qy   2669 TATTAGACACCAACACAGA---------------------------------------AA 2649
Db    354 yGlnArgGlyAlaHisGlyMetProGlyLysProGlyProMetGlyProLeuGlyIlePr 374

Qy   2648 AGCTAGCAATGGATTCCCTTCTACTTTGTTAAATAAATAAGTTAAATATTTAAATGCCTG 2589
Db    374 oGlySerSerGlyPheProGlyAsn---------------------------------- 382

Qy   2588 TGTCTCTGTGATGGCAACAGAAGGACCAACAGGCCACATCCTGATAAAAGGTAAGAGGGG 2529
Db    383 -ProGlyMetLysGlyGluAlaGlyProThrGly---------AlaArgGlyProGluGl 399

Qy   2528 GGTGGATCAGCAAAAAGACAGTGCTGTGGGCTGAGGGGACCTGGTTCTTGTGTGTTGCCC 2469
Db    399 yProGlnGlyGlnArgGlyGluThrGlyProProGlyProValGlySerProGlyLeuPr 419

Qy   2468 CTCAGGACTCTTCCCCTACAAATAAGTCATATGTTCAAATCCCATGGAGGAGTGTTTCAT 2409
Db    419 oGlyAlaIleGly---------------------------------------------- 423

Qy   2408 CCTAGAAACTCCCATGCAAGAGCTACATTAAACGAAGCTGCAGGTTAAGGGGCTTAGAGA 2349
Db    424 --------------------------------------------------------ThrAs 425

Qy   2348 TGGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCTCA 2289
Db    425 pGlyThrProGly-----------------ProLysGlyProThr-------------- 434

Qy   2288 ACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCAGAGTCCCCGCATTC 2229
Db    435 -----------------------------GlySerProGlyThrSerGlyProProGl 444

Qy   2228 CAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGG 2169
Db    444 ySerAla---GlyProProGlySerProGly---ProGlnGlySerThrGlyProGlnGl 462

Qy   2168 CCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGG---------------AGAGGGAC 2124
Db    462 y---AsnSerGlyLeuProGlyAspProGlyPheLysGlyGluAlaGlyProLysGlyGl 481

Qy   2123 GCCCCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGAGCC 2064
Db    481 uProGlyProHisGlyIleGln-------------------------------------- 488

Qy   2063 ACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAACAG 2004
Db    489 ---------------------------GlyProIleGlyProProGlyGluGluGlyLy 499

Qy   2003 GAGCGGGGAGCTGGGA---CCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTT 1947
Db    499 sArgGlyProArgGlyAspProGlyThrLeuGlyProPro------------------- 512

Qy   1946 TCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAA 1887
Db    513 ---------------GlyProVal---------------------------------- 515

Qy   1886 TGGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTG-------- 1835
Db    516 -GlyGluArgGlyAlaProGlyAsnArgGlyPheProGlySerAspGlyLeuProGlyPr 535

Qy   1834 -------GCTGAGCTGGACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCA 1782
Db    535 oLysGlyAlaGlnGlyGluArgGly-ProValGly-SerSerGlyProLysGlySerGln 554

Qy   1781 GGAAGGCACTATCCAGGATGGCGAGGTCCAGGCAGATGCCCCGGCCCGGAACCACCCTGG 1722
Db    555 Gly--------------------AspProGlyArgProGlyGluPro-GlyLeuProGl 567

Qy   1721 CCTCGGTGGGCTCACCCACCACCACACGTACGGAGACATCACAGGCAGA---GGCCCCGC 1665
Db    567 yAlaArgGlyLeuThrGlyAsnProGlyValGlnGlyProGluGlyLysLeuGlyProLe 587

Qy   1664 AGAGCGCGGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGGG- 1606
Db    587 u-GlyAlaProGlyGluAspGlyArgProGlyProProGly---------SerIleGlyI 604

Qy   1605 --AAGGGAGCTCCAGGC---TTAGGG---CCTGGCAGGAAGCTGGTCATCAGGCTGTCCT 1554
Db    604 leLysGlyGlnProGlyThrMetGlyLeuProGlyProLys-GlySerAsnGlyAspPro 623

Qy   1553 CACTGCTAGCACCTCCAGTGTCCCCTCGGTATT------TGGGCAGGAACACCTGCTTCT 1500
Db    624 --------------------------GlyLysProGlyGluAlaGlyAsnProGlyVal 634

Qy   1499 CCCGGTGGTAGAGGGAGG-----------------------CCAGTGTGTAGGGCAGGA 1464
Db    635 ProGlyGlnArgGlyAlaProGlyLysAspGlyLysValGlyProTyrGlyProProGly 654

Qy   1463 TCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACAC 1404
Db    655 ProProGlyLeuArg-GlyGluArgGlyGluGlnGly--ProProGlyProThrGlyPhe 673

Qy   1403 TGTGGGACAGGCATGTGGCACCGGCAGCCACAGGGAAAG--------------------- 1365
Db    674 GlnGly-----------HisProGlyProProGlyProProGlyGluGlyGlyLysPro 689

Qy   1364 -----------CTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCTGCACCA 1317
Db    690 GlyAspGlnGlyValProGlyGlyProGlyAlaVal-----------------GlyPro 703

Qy   1316 GCCGGTCCATGACCAGAGAGA---AGACCAGGGAGATGGCGCACTGCAGGAACAGCCCCA 1260
Db    704 LeuGlyProArgGlyGluArgGlyAsnProGlyGluArgGlyGluProGlyIleThr--- 722

Qy   1259 GGCTGCCCATCCGAACGC---------CTTCATCATAGTGTCTCCGGGCCTCGG------ 1215
Db    723 GlyLeuProGlyGluLysGlyMetAlaGlyGlyHisGlyProAspGlyProLysGlySer 742

Qy   1214 -----------------------------TGCCCGGCTCAGCTCTGGGCACGCCCTGGT 1185
Db    743 ProGlyProSerGlyThrProGlyAspThrGlyProProGlyLeuGlnGlyMetProGly 762

Qy   1184 ACA--------------GCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCA 1140
Db    763 GluArgGlyIleAlaGlyThrProGlyProLysGlyAsp-------------------- 775

Qy   1139 TGAGTGCCATCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGTGCGGGGCATGC 1080
Db    776 --------------------------------ArgGlyGlyIleGlyGluLysGlyAla 784

Qy   1079 GGCAGCACAGCTGGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGG 1020
Db    785 GluGlyThrAlaGlyAsnAspGlyAlaGlyGlyLeuProGly---------ProLeuGly 801

Qy   1019 CCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGG 960
```

```
                 |||                            ||||||     ||||||
Db        802 Pro-------------------------ProGlyProAlaGlyLeuLeu-------- 809

Qy        959 TGGGGCCCAGCGCTGCCTCCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGA 900
                 |||           |||   |||                        |||
Db        810 ---GlyGluLysGlyGluProGlyProArg-----------------------GlyLeu 820

Qy        899 TGAGGGTGAGCAGGCCAAAGAGGCACTCCT------------------------------ 870
                 |||     |||   ||||||   |||
Db        821 ValGlyProProGlySerArgGlyAsnProGlySerArgGlyGluAsnGlyProThrGly 840

Qy        869 -------------------------------------CCTGGGTGCCCAGGTAGGGGG 849
                                                   ||||||      |||
Db        841 AlaValGlyPheAlaGlyProGlnGlySerAspGlyGlnProGlyValLysGlyGlu--- 859

Qy        848 CCAGGGCACTGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGAC 789
              ||||||        |||   |||          |||||||||||    ||||||
Db        860 ProGlyGlu------ProGlyGlnLysGlyAspAlaGlySerProGly---ProGlnGly 876

Qy        788 TGATCATGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGTCCGGGT------------ 741
                                         ||||||     ||||||
Db        877 ---------------LeuAlaGlySerProGlyProHisGlyProAsnGlyValProGly 891

Qy        740 ------------------------------------------------CCCGGAAGAGGT 729
                                                               ||||||
Db        892 LeuLysGlyGlyArgGlyThrGlnGlyProProGlyAlaThrGlyPheProGlySerAla 911

Qy        728 CAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCA 669
                 |||   |||||||||              ||||||        ||||||   |||
Db        912 GlyArgValGlyProProGlyProAlaGlyAlaProGly---------ProAlaGlyPro 928

Qy        668 ------CGCCCAGGATGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGCACAGCAGCC 615
                    ||||||         |||         |||||||||:::
Db        929 LeuGlyGluProGlyLysGluGlyProProGlyProArgGlyAspProGlySerHisGly 948

Qy        614 CTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCAAGGACAGTG 555
                 :::                            ||||||   ||||||   |||
Db        949 ArgValGly-----------------ValArgGlyProAlaGlyProProGly------ 960

Qy        554 CCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTA 495

Db        960 ------------------------------------------------------------ 960

Qy        494 GGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCATGGTCATGAACT 435
              |||   |||   :::      |||      |||      |||   |||
Db        961 GlyProGlyAspLysGlyAspProGlyGluAspGlyGlnProGlyProAspGly------ 978

Qy        434 TCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCCAAAC 375
                        ||||||   ||||||            |||      :::
Db        979 --------ProProGlyProAlaGlyThrThrGlyGlnArgGlyIleVal--------- 992

Qy        374 ACACCTCCAGGCCAAAGG----------------------------TTAGCAGGTTGA 345
                    ||||||||||||                              ||||||
Db        993 GlyMetProGlyGlnArgGlyGluArgGlyMetProGlyLeuProGlyProAlaGlyThr 1012

Qy        344 CCAGCAAGAGCTGGGCTTTCCGGTGCCGCAGCAGGC--------------------GGC 306
              |||   :::   |||   :::||||||                            |||
Db       1013 ProGlyLysValGlyProThrGlyAlaThrGlyAspLysGlyProProGlyProValGly 1032

Qy        305 TCACCCACAGCCTCTGGACCATAGTGGG---CCAGGCGGGTAGGGCTCAGGGGGCCGTTC 249
                 |||   ::::::||||||   |||   ||||||      |||      |||
Db       1033 ProProGlySerAsnGlyPro-ValGlyGluProGlyProGluGlyProAlaGlyAsnAs 1052

Qy        248 AGGCACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTG 189
              ||||||||            |||   |||:::      |||
Db       1052 pGlyThrProGly------ArgAspGlyAlaValGlyGluArgGly------------- 1065

Qy        188 CTGCCGCCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGG 129
                                  |||   ||||||                  |||   |||
Db       1066 ------------AspArgGlyAspProGly--------------ProAlaGlyLeuPr 1076

Qy        128 TCCAGCTTCTCAGCCCATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGT 69
                 :::                  ||||||     ||||||        ||||||
```

```
Db       1076 oGlySerGln-GlyAlaProGlyThrProGlyProValGly---AlaProGlyAspAlaG 1095

Qy         68 CTCATCACTCAGATCCTGGC----------------------------CGAGGCGCGC 39
                            |||||||||                            |||   |
Db       1095 lyGlnArgGlyAspProGlySerArgGlyProIleGlyHisLeuGlyArgAlaGlyLysA 1115

Qy         38 GGCTGTCACCCGGAGCCAGCGCGTGCAGG 10
                ||    ||||||         |||
Db       1115 rgGlyLeuProGlyProGlnGlyProArg 1124
```

RESULT 24
CGHU2E
collagen alpha 2(XI) chain precursor - human (fragment)
N;Alternate names: procollagen alpha 2(XI) chain
N;Contains: proline/arginine-rich protein (PARP)
C;Species: Homo sapiens (man)
C;Date: 07-Jun-1990 #sequence_revision 03-Oct-1995 #text_change 22-Jun-1999
C;Accession: S34790; A32645
R;Zhidkova, N.I.; Brewton, R.G.; Mayne, R.
FEBS Lett. 326, 25-28, 1993
A;Title: Molecular cloning of PARP (proline/arginine-rich protein) from human cartila
hain.
A;Reference number: S34790; MUID:93314796; PMID:8325374
A;Accession: S34790
A;Molecule type: mRNA
A;Residues: 1-663 <ZHI>
A;Cross-references: EMBL:L18987; NID:g306439; PIDN:AAA35498.1; PID:g306440
R;Kimura, T.; Cheah, K.S.E.; Chan, S.D.H.; Lui, V.C.H.; Mattei, M.G.; van der Rest, M
J. Biol. Chem. 264, 13910-13916, 1989
A;Title: The human alpha2(XI) collagen (COL11A2) chain. Molecular cloning of cDNA and
A;Reference number: A32645; MUID:89340485; PMID:2760050
A;Accession: A32645
A;Molecule type: DNA; mRNA
A;Residues: 586-1546 <KIM>
A;Cross-references: GB:J04974; NID:g180714; PIDN:AAA52034.1; PID:g180715
A;Note: parts of this sequence were determined by protein sequencing
C;Comment: Prolines and lysines at the third position of the tripeptide repeating uni
ed and subsequently O-glycosylated.
C;Genetics:
A;Gene: GDB:COL11A2
A;Cross-references: GDB:119788; OMIM:120290
A;Map position: 6p21.3-6p21.3
A;Introns: 1302/3; 1320/3; 1332/3; 1350/3; 1440/1; 1477/3
A;Note: the list of introns is incomplete
C;Complex: type XI collagen may be a heterotrimer of two alpha 1(XI) chains (see PIR:
3(XI) chain (see PIR:CGHU6C), initially linked by disulfide bonds among their carbox
rmed with desmosine cross-links made from lysine and allysine residues
C;Function:
A;Description: structural component of extracellular fibrous polymer associated with
A;Note: may play a role in controlling the lateral growth of collagen II fibrils
C;Superfamily: collagen alpha 1(V) chain; fibrillar collagen carboxyl-terminal homolo
C;Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxypr
F;1-254/Domain: non-collagenous (fragment) #status predicted <NC3>
F;1-187/Product: proline/arginine-rich PARP protein (fragment) #status predicted <PAR
F;255-305/Domain: collagenous, triple helix #status predicted <COL2>
F;306-342/Domain: non-collagenous #status predicted <NC2>
F;343-1356/Region: helical
F;429-431/Region: cell attachment (R-G-D) motif
F;447-449/Region: cell attachment (R-G-D) motif
F;1257-1259/Region: cell attachment (R-G-D) motif
F;1357-1380/Region: carboxyl-terminal nonhelical telopeptide
F;1381-1546/Domain: carboxyl-terminal propeptide (fragment) #status predicted <CTP>
F;1403-1546/Domain: fibrillar collagen carboxyl-terminal homology (fragment) #status
F;109-163,1511-1545/Disulfide bonds: #status predicted
F;319/Modified site: allysine (Lys) #status predicted
F;426,1266/Modified site: 5-hydroxylysine (Lys) #status predicted
F;426,1266/Binding site: carbohydrate (Lys) (covalent) #status predicted
F;927,933,1008,1017,1035,1038,1290,1296,1305,1317,1320/Modified site: 4-hydroxyprolin
F;929/Modified site: 4-hydroxyproline (Pro) #status atypical
F;942,1023,1299/Modified site: 5-hydroxylysine (Lys) #status experimental
F;942,1023,1299/Binding site: carbohydrate (Lys) (covalent) #status experimental
F;1427,1433,1450,1459/Disulfide bonds: interchain #status predicted
F;1460/Binding site: carbohydrate (Asn) (covalent) #status predicted

```
Alignment Scores:
Pred. No.:              4.7e-12      Length:         1546
Score:                  324.50       Matches:        302
Percent Similarity:     28.85%       Conservative:   58
Best Local Similarity:  24.20%       Mismatches:     409
Query Match:            5.21%        Indels:         479
DB:                     1            Gaps:           70

US-09-759-143-110 (1-3410) x CGHU2E (1-1546)

Qy     3224 TTATTGCAAACGGCACTTAAACCCCCCCTGAGAGATAAGACCTCCCTTAGCTCAGGCAGG 3165
            ::::|||::::::::|||   |||||||||    :::
Db      218 IleLeuGluSerSerLeuLeuProProLeuGluGluAlaAla------------------ 231

Qy     3164 GGGTGCTCCTGAGTTTCTGTGTGAGATTCCCCAAGCACAGATATACTCTGGGGGCTGAGA 3105
                                     |||                |||||||:::
Db      232 ---------------------HisGlyProArg------------------GlyLeuLys 238

Qy     3104 TGGACAAAG------------GCTTGGGAAACCGCACTTTGTGCTTCTGGTCCTGCAGTA 3057
               |||          |||      :::        ||||||
Db      239 GlyGluLysGlyGluProAlaValLeuGluProGlyMetLeuValGluGlyProPro--- 257

Qy     3056 GCTCCAAACAGGGTTGTGGAGCTGGTGGGGAAAGTTGGGGGTAGGGGAAAGTTGGGGGTA 2997
               |||        |||:::|||      |||    :::|||       |||
Db      258 GlyProGluGlyProAlaGlyLeuIleGlyProProGlyIleGlnGlyAsnProGly--- 276

Qy     2996 GGGGAAATTTTGGGCAGTGCCTTCATCAGCCCAGTCCTAGAGAGAGTAGAGGGGAGTGGA 2937
                                            ||||||                   |||
Db      277 -------------------------------ProVal--------------------Gly 279

Qy     2936 AGTGGGGGGAACCAGGCTGGGCCAAGAGAAGAGGGGTGGTTAGGGAAGCCGTTGAGACCT 2877
                |||    :::       |||          |||      |||         |||
Db      280 AspProGlyGluArgGlyProProGlyArgAlaGlyLeuProGlySerAspGlyAlaPro 299

Qy     2876 GAAGCCCCACCCTCTACCTTCCTTCAACACCCTAACCTTGGGTAACAGCATTTGGAATTA 2817
                |||||     |||   |||     |||
Db      300 ---GlyProProGlyThrSerLeuMetLeuPro---------------------------Phe 310

Qy     2816 TCATTTGGGATGAGTAGAATTTCCAAGGTCCTGGGTTAGGCATTTTGGGGG------GCC 2763
               ||||||          |||                     |||
Db      311 ArgPheGlySerGlyGlyGlyAspLys---------------------GlyProValVal 323

Qy     2762 AGACCCCAGGAGAAGAAGATTCTGGCAATGATCAGCCCAATGACCAGCTATCTCAGGGGA 2703
                 ||||||     :::      |||::::::          |||||||||
Db      324 AlaAlaGlnGluAlaGlnAlaGlnAlaIleLeuGlnGlnAlaArgLeuAlaLeuArgGly 343

Qy     2702 CCTGATTGTTGGGGATCCCCCACCCTACCCAAATATTAGACACCAACACAGAAAAGCTAG 2643
            |||                 |||           |||
Db      344 Pro---------------ProGlyProMetGlyTyr------------------------ 350

Qy     2642 CAATGGATTCCCTTCTACTTTGTTAAATAAATAAGTTAAATATTTAAATGCCTGTGTCTC 2583

Db      350 ------------------------------------------------------------ 350

Qy     2582 TGTGATGGCAACAGAAGGACCAACAGGCCACATCCTGATAAAAGGTAAGAGGGGGGTGGA 2523
                                 |||   ||||||   |||
Db      351 -----------------ThrGlyArgProGlyPro------------------------- 356

Qy     2522 TCAGCAAAAAGACAGTGCTGTGGGCTGAGGGGACCTGGTTCTTGTGTGTTGCCCCTCAGG 2463
                                     |||      |||||||||     |||
Db      357 ------------------------LeuGlyGlnProGlySerProGlyLeu--------- 365

Qy     2462 ACTCTTCCCCTACAAATAAGTCATATGTTCAAATCCCATGGAGGAGTGTTTCATCCTAGA 2403
                                          |||          |||      |||:::
Db      366 ------------------------------LysGlyGluSerGlyAspLeuGlyProGln 375

Qy     2402 AACTCCCAT--------------------------------------------------- 2394

Db      376 GlyProArgGlyProGlnGlyLeuThrGlySerLeuGlyLysAlaGlyArgArgGlyArg 395

Qy     2393 ---------------GCAAGAGCTACATTAAACGAAGCTGCAGGTTAAGGGGCTTAGA-- 2351
                            ||||||   ||||||    :::        |||||||||   |||
Db      396 AlaGlyProAspGlyAlaArgGlyThrLeuGlyAspProGly-ValLysGlyAspArgGl 415

Qy     2350 ----GATGGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGT 2295
                 ||||||  ||||||
Db      415 yPheAspGlyLeuProGly----------------------------------------- 421

Qy     2294 GTCTCAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCAGAGTCCCC 2235
                                                       |||   |||
Db      422 -------------------------------------------LeuProGlyGluLysGl 427

Qy     2234 GCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACCCCCTT 2175
             |||            |||   |||||||||
Db      427 yHisArgGlyAspThrGlyProArgGlyLeuPro-------------------------- 438

Qy     2174 GGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGG----------------- 2132
                             |||   |||   :::   |||||||||
Db      439 ----------------GlyProProGlyGluAspGlyGluArgGlyAspAspGlyGluIl 453

Qy     2131 -------AGAGGGACGCCCCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGG 2079
                   ||||||   |||                       |||      :::||
Db      453 eGlyProArgGlyLeuPro---------------------GlyGluSerGlyProArgGl 466

Qy     2078 T------GGCAGCAGAGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCG 2025
            |      |||   :::   |||                         |||
Db      466 yLeuLeuGlyProLysGlyProProGlyIleProGlyProProGlyValArgGlyMetAs 486

Qy     2024 GCAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGGACCC---AGTGAGGCAGGCCCTCC 1968
               |||   |||             |||   |||||||||     |||   |||||||
Db      486 pGlyProGlnGlyPro------LysGlySerLeuGlyProGlnGlyGluProGlyProPr 504

Qy     1967 ACCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATA 1908
            |  |||           |||      :::   |||                     |||
Db      504 oGlyGlnGlnGlyThrPro--GlyThrGlnGlyLeu---ProGlyProGlnGlyAlaIle 522

Qy     1907 CTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGG 1848
               |||      :::         :::   |||||||||            |||
Db      523 GlyProHisGlyGluLysGlyProGlnGlyLysProGlyLeuProGlyMetProGlySer 542

Qy     1847 CAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGG 1788
                  |||      |||:::                     ||||||||||||
Db      543 AspGlyProProGlyHisProGlyLys------------------GluGlyProProGly 556

Qy     1787 ACAGCAGGAAGGCACTATCCA---GGATGGCGAGGTCCAGGCAGATGCCCCGGCCCGGAA 1731
            |||   |||:::      |||   |||   :::||||||        |||||||||:::
Db      557 ThrLysGlyLysProGlyProSerGlyProGlnGlyProLeuGlyTyrProGlyProGln 576

Qy     1730 CCAC-----------CCTGGCCTCGGTGGGCTCACCCACCACCAC--------------- 1697
                              |||:::   ||||||      |||
Db      577 -GlyValLysGlyValAsnGlyIleArgGlyLeuLysGlyHisLysGlyGluLysGlyGl 596

Qy     1696 ----------------ACGTACGGAGACATCACAGGCAGAGGCCCC---GCAGAGCGCGG 1656
                               |||||||||      :::|||         |||   |
Db      596 uAspGlyPheProGlyPheLysGlyAspIleGlyValLysGlyAspArgGlyGlu-ValG 616

Qy     1655 GTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGGGAAGGGAGCTC 1596
            ||      |||||||||                  |||      |||   |||:::
Db      616 lyValProGlySerArgGly---GluAspGlyProGluGlyProLysGlyArgThrGlyP 635

Qy     1595 CAGGCTTAGGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCTAGCACCTCCAG 1536
            ||         ||||||          ::::::             |||      |||
Db      635 roThrGlyAspProGlyProProGlyLeuMetGlyGluLysGlyLysLeuGlyValProG 655

Qy     1535 TGTCCCCTCGGTATTTGGGCAGGAACA------------------------CCTGCTTCT 1500
                  |||   |||   ||||||                            |||
Db      655 lyLeuProGlyTyrProGlyArg-GlnGlyProLysGlySerLeuGlyPheProGlyPhe 674

Qy     1499 CCCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACC 1440
            ||||||      |||                  |||:::   ||||||
Db      675 ProGlyAlaSerGly---------------GluLysGlyAlaArgGlyLeuSerGly------ 688

Qy     1439 CGGTGAGGGCGGCTGAAGCTGTCACCACGG-------------CCACACTGTGGACAGGC 1392
                           |||      ||||||              ||||||      ||||||
Db      689 ---------------LysSerGlyProArgGlyGluArgGlyProHisGlySerThrGly 703
```

```
Qy   1391 ATGTGGCACCGGCAGCCACAGGGAAAGCTG------------------------------ 1362
                   |||   |||:::|||
Db    704 ------GlnArgGlyProArgGlyAlaThrGlyLysSerGlyAlaLysGlyThrSerGly 721

Qy   1361 ---------CCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCTGCACCAGCCGGT 1311
                   ||||||   |||              |||      :::      |||:::|||
Db    722 GlyAspGlyProHisGlyProProGly------GluArgGlyLeuProGlyProSerGly 739

Qy   1310 ------------CCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACAGCC 1263
                      |||   |||:::        ||||||               |||
Db    740 SerAsnGlyPheProGlyProLysGlyProProGlyProProGlyLysAspGly---Leu 758

Qy   1262 CCAGGCTGCCCATCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAG 1203
          ||||||  |||    :::|||                 |||          |||
Db    759 ProGlyHisProGlyGlnArgGlyGluValGlyPheGlnGlyLys***GlyPro------ 776

Qy   1202 CTCTGGGCACGCCCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGG 1143
                      ||||||
Db    777 ------------ProGly------------------------------------------ 778

Qy   1142 TCATGAGTGCCATCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCA 1083
                                        |||                     ||||||
Db    779 ------------------------------ProProGlyValValGlyProGlnGlyAla 788

Qy   1082 TGCGGCAGCACAGCTGGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGC 1023
            |||   |||    |||         ||||||   |||   ||||||            |||
Db    789 AlaGlyGluThr---GlyProMetGlyGluArgGlyHisProGlyProProGlyProPro 807

Qy   1022 GGG------------CCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCC 975
          |||            ||||||                 |||:::        |||
Db    808 GlyGluGlnGlyLeuProGlyThrAlaGlyLysGluGlyThrLysGlyAspProGly--- 826

Qy    974 CTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCTCCTCAGCCACCAGCAGTGTGGCTGCTA 915
                                        ||||||      |||
Db    827 ------------------------------ProProGlyAlaPro--------------- 831

Qy    914 CGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGCACTCCT--------------- 870
                   |||:::   |||   ||||||   ||||||   |||
Db    832 ---------GlyLysAspGlyProAlaGlyLeuArgGlyPheProGlyGluArgGlyLeu 848

Qy    869 CCTGGGTGCCCAGGTAGGGGGCCAGGG--------------------------------- 843
          ||||||      |||   |||||||||
Db    849 ProGlyThrAlaGly---GlyProGlyLeuLysGlyAsnGluGlyProSerGlyProPro 867

Qy    842 ------CACTGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGAC 789
                         |||   ::       |||   ||||||   |||   |||
Db    868 GlyProAlaGlySerProGlyGluArgGlyAlaAlaGlySer---GlyGlyProIleGly 886

Qy    788 TGATCATGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGTCCGGGTCCCGGAAGAGGT 729
                         ||||||   |||||||||      ||||||               |||
Db    887 ---------------ArgGlnGlyArgProGlyProGlnGlyProProGlyAlaAlaGly 901

Qy    728 CAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCA 669
          ::::::                           ||||||   :::   |||      |||
Db    902 GluLys---------------------GlyValProGlyGluLysGlyProIleGlyPro 914

Qy    668 CGC---------------------------------CCAGGATGAGCAGTGCCAGCTCCA 642
                                              ||||||   |||   |||   |||
Db    915 ThrGlyArgAspGlyValGlnGlyProValGlyLeuProGlyProAlaGlyPro---Pro 933

Qy    641 GGGGCC------------TGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGG------ 600
          |||                  |||   |||            :::         :::
Db    934 GlyValAlaGlyGluAspGlyAspLysGlyGluValGlyValProGlyGlnLysGlyThr 953

Qy    599 ---------------------------------------------------CCCTTGGGA 591
                                                             |||:::|||
Db    954 LysGlyAsnLysGlyGluHisGlyProProGlyProProGlyProIleGlyProValGly 973

Qy    590 TGAGAAAGAGGCTCAGCAGGATGCCCAAGGACAGTGCCCAGATGAAGGGCCGGCGGCGGC 531
             :::   |||:::||||||                        |||   ||||||
Db    974 ---GlnProGlyAlaAlaGly------------------------ArgAspGlyGlyAla 984

Qy    530 CATAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTAGGA---GCGGGACACAGA------ 480
                :::         ||||||:::      :::      |||      |||   |||
Db    985 GlyGlyLeuGlyThrGlnGlyHisPheGlyAlaLysGlyAspGluGlyThrArgGlyPhe 1004

Qy    479 ---------CCAGGCCCA---GCACTGGACCAATGCCCAGCACCATGGTCATGAACTTCT 432
                   |||||||||               |||         |||   |||
Db   1005 AsnGlyProProGlyProIleGlyLeuGlnGlyLeuProGlyProSerGlyGluLysGly 1024

Qy    431 ------------------------------CCTCTACCCCCACTTCCAGCAGCAGAGGCG 402
                                        |||   ||||||   |||         |||
Db   1025 GluThrGlyAspValGlyProMetGlyProProGlyProProGlyProArgGlyProAla 1044

Qy    401 GCACATAGGTGATGCCTGCGGCCAAAC--------------------------------- 375
                               ||||||
Db   1045 Gly------------------ProAsnGlyAlaAspGlyProGlnGlyProProGlyGly 1058

Qy    374 ------------ACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTT 327
                         |||||||||::::::      |||   |||      :::|||
Db   1059 ValGlyAsnLeuGlyProProGlyGluLys------GlyGluProGlyGluSerGlySer 1076

Qy    326 TCCGGTGCCGCAGCAGGCGGCTCACCCACAGCCTCTGGACCATAGTGGGCCAGGCGGGTA 267
             |||            |||   |||         ||||||      |||
Db   1077 ProGlyIleGln------GlyGluProGlyValLysGlyPro-ArgGlyGluArg----- 1092

Qy    266 GGGCTCAGGGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCAGAAGC 207
           |||      |||   ||||||   |||                 |||
Db   1093 -GlyGluLysGlyGluSerGlyGlnPro--------------GlyGluProGlyProPr 1107

Qy    206 TGCGGCCTCTCCTCCTTGCTGCCGCCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCT- 148
              |||      |||                 |||:::    ||||||  |||:::
Db   1107 oGlyAlaLysGlyProGlnGlyAspAspGlyProLysGlyAsnProGly-ProValGlyP 1127

Qy    147 -----GCCAGCCCTTTGGTGCCGGTCCAGCTTCTCAGCCCATGCTCAACACCTGCTGCTG 93
                     |||      |||   :::                       |||
Db   1127 heProGlyAspProGlyProProGlyGlu--------------GlyGlyProArgGlyG 1142

Qy     92 TGGGGCACCTCAGTGGGGACACGTCTCATCACTCAGATCCTGGCCGAGGCGCGCGGCTGT 33
                        ||||||                   :::||||||:::            |
Db   1142 lnAspGlyAlaLysGlyAspArgGlyGluAspGlyGluProGlyGlnProGly------S 1160

Qy     32 CACCCGGA 25
          ||||||||
Db   1160 erProGly 1162
```

RESULT 25
S38196
sucrose transport protein SUC2 - Arabidopsis thaliana
N;Alternate names: sucrose-proton symporter SUC2
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 01-Dec-2000
C;Accession: S38196; T00773
R;Sauer, N.K.
submitted to the EMBL Data Library, October 1993
A;Description: SUC1 and SUC2: two sucrose transporters from Arabidopsis thaliana.
A;Reference number: S38196
A;Accession: S38196
A;Molecule type: mRNA
A;Residues: 1-512 <SAU>
A;Cross-references: EMBL:X75382; NID:g407091; PID:g407092
R;Vysotskaia, V.S.; Schwartz, J.R.; Toriumi, M.; Yu, G.; Oji, O.; Kwan, A.; Liu, S.;
rtz, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federspiel
submitted to the EMBL Data Library, July 1998
A;Description: Arabidopsis thaliana chromosome 1 BAC T22J18 sequence, complete sequen
A;Reference number: Z14202
A;Accession: T00773
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-490,'A',492-512 <VYS>
A;Cross-references: EMBL:AC003979; NID:g3172156; PID:g3287687; GSPDB:GN00059; ATSP:T2
A;Experimental source: cultivar Columbia
C;Genetics:
A;Gene: SUC2; ATSP:T22J18.12
A;Map position: 1

```
A;Introns: 419/3; 441/1; 455/3
C;Superfamily: common tobacco sucrose transport protein
C;Keywords: sugar transport

Alignment Scores:
Pred. No.:              5.4e-12       Length:        512
Score:                  323.50        Matches:       123
Percent Similarity:     42.07%        Conservative:  92
Best Local Similarity:  24.07%        Mismatches:    195
Query Match:            5.04%         Indels:        101
DB:                     2             Gaps:          18

US-09-759-143-110 (1-3410) x S38196 (1-512)

Qy      323 CGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCC 382
             |||     :::::::::  |||:::   :::      |||:::::::        |||
Db       28 ArgLeuArgLysIleIleSerValSerSerIleAlaAlaGlyValGlnPheGlyTrpAla 47

Qy      383 GCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATG 442
                  ::::::  :::  |||  :::        :::|||:::        |||:::
Db       48 LeuGlnLeuSerLeuLeuThrProTyrValGlnLeuLeuGlyIleProHisLysTrpAla 67

Qy      443 ACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCA 502
            :::::::::           |||||||:::    |||::::::      |||::::::|||
Db       68 SerLeuIleTrpLeuCysGlyProIleSerGlyMetLeuValGlnProIleValGlyTyr 87

Qy      503 GCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCC 562
               ||||||           |||:::||||||||||||||||||||   |||    :::
Db       88 HisSerAspArgCysThrSerArgPheGlyArgArgArgProPheIleValAlaGlyAla 107

Qy      563 TTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTG--- 619
                :::   ::::::::|||||||||         |||:::   |||   :::
Db      108 GlyLeuValThrValAlaValPheLeuIle---------GlyTyrAlaAlaAspIleGly 124

Qy      620 ------------------CTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATC 661
                              |||      |||:::        :::|||:::
Db      125 HisSerMetGlyAspGlnLeuAspLysProProLysThrArgAlaIleAlaIlePheAla 144

Qy      662 CTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTG 721
             ||||||      :::||||||                   |||      |||
Db      145 LeuGlyPheTrpIleLeuAspValAlaAsnAsnThrLeuGlnGlyProCysArgAlaPhe 164

Qy      722 CTCTCTGACCTCTTCCGG---GACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCC 778
            |||:::||||||         :::            |||   |||   :::   ::::::
Db      165 LeuAlaAspLeuSerAlaGlyAsnAlaLysLysThrArgThrAlaAsnAlaPhePheSer 184

Qy      779 TTCATGATCAGTCTTGGGGGCTGCCTGGGCTAC--------------------CTCCTG 817
             |||   ::::::::|||      ||||||||                      |||
Db      185 PhePheMetAlaValGlyAsnValLeuGlyTyrAlaAlaGlySerTyrArgAsnLeuTyr 204

Qy      818 CCTGCCATTGACTGGGAC---ACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 874
                :::   :::      |||   :::         |||               :::
Db      205 LysValValProPheThrMetThrGluSerCysAspLeuTyrCysAlaAsnLeuLysThr 224

Qy      875 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 934
             |||   |||      :::||||||:::   |||               :::|||
Db      225 CysPhePheLeuSerIleThrLeuLeu---LeuIleValThrPheValSerLeuCysTyr 243

Qy      935 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 994
               :::|||         |||      |||:::|||   :::
Db      244 ValLysGluLysProTrpThrProGluProThrAlaAspGlyLysAla------------ 259

Qy      995 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCCTGCTT 1054
                                   :::   :::   |||         |||   :::
Db      260 -----------------------SerAsnValProPhe------PheGlyGluIlePhe 269

Qy     1055 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1114
                        :::|||             |||   :::      |||   :::
Db      270 GlyAlaPheLysGluLeu------------LysArgProMetTrpMetLeuLeuIleVal 285

Qy     1115 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 1174
                   :::|||:::|||         |||   ||||||   ||||||:::::::|||
Db      286 ThrAlaLeuAsnTrpIleAlaTrpPheProPheLeuLeuPheAspThrAspTrpMetGly 305

Qy     1175 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 1234
                :::|||   |||.         |||                  :::::::::  |||
Db      306 ArgGluValTyrGlyGlyAsnSerAspAlaThrAlaThrAlaAlaSerLysLysLeuTyr 325

Qy     1235 GATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTC 1294
            ::::::|||||||||      |||:::|||||||||   |||          :::
Db      326 AsnAspGlyValArgAlaGlyAlaLeuGlyLeuMetLeuAsnAlaIleValLeuGlyPhe 345

Qy     1295 TTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCC 1354
               ||||||   ::::::   :::   ::::::   |||      |||      |||
Db      346 MetSerLeuGlyValGluTrpIleGlyArgLysLeuGly---GlyAlaLysArgLeuTrp 364

Qy     1355 AGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTG 1414
               :::      |||   :::                |||   ||||||:::   :::|||   |||
Db      365 GlyIleValAsnPheIleLeu---------AlaIleCysLeuAlaMetThrValValVal 381

Qy     1414 ------------------------------------------------------------ 1414

Db      382 ThrLysGlnAlaGluAsnHisArgArgAspHisGlyGlyAlaLysThrGlyProProGly 401

Qy     1415 ---GTGACAGCTTCAGCC------------GCCCTCACCGGG------------TTCACC 1447
               |||||||||   |||               |||:::   |||               |||
Db      402 AsnValThrAlaGlyAlaLeuThrLeuPheAlaIleLeuGlyIleProGlnAlaIleThr 421

Qy     1448 TTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAG--- 1504
             ||||||            :::|||:::   |||||||||::::::
Db      422 PheSer------------IleProPheAlaLeuAlaSerIlePheSerThrAsnSerGly 437

Qy     1505 -----------------------------------CAGGTGTTCCTGCCCAAATACCGA 1528
                                                :::   :::|||:::
Db      438 AlaGlyGlnGlyLeuSerLeuGlyValLeuAsnLeuAlaIleValValProGlnMetVal 457

Qy     1529 GGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCT 1588
                     ||||||         ::::::
Db      458 IleSerValGlyGlyGlyProPheAspGluLeuPheGlyGlyGlyAsnIleProAla--- 476

Qy     1589 AAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTC--- 1645
                              |||   |||   :::   |||      ||||||:::|||
Db      477 ---------------PheValLeuGlyAlaIleAlaAlaAlaValSerGlyValLeuGly 491

Qy     1646 ------------------CCACCTCCACCCGCG 1660
                              |||||||||   |||
Db      492 LeuThrValLeuProSerProProProAspAla 502
```

```
RESULT 26
G86360
probable sucrose-proton symporter SUC2 protein T22J18.12 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Nov-2001
C;Accession: G86360
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: G86360
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-512 <STO>
A;Cross-references: GB:AE005172; NID:g3287687; PIDN:AAC25515.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1
C;Superfamily: common tobacco sucrose transport protein

Alignment Scores:
Pred. No.:              5.4e-12       Length:        512
```

```
Score:                    323.50     Matches:        123
Percent Similarity:       42.07%     Conservative:   92
Best Local Similarity:    24.07%     Mismatches:     195
Query Match:              5.04%      Indels:         101
DB:                       2          Gaps:           18

US-09-759-143-110 (1-3410) x G86360 (1-512)

Qy        323 CGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCC 382
                 |||     :::::::::   |||:::    :::       |||:::::::          |||
Db         28 ArgLeuArgLysIleIleSerValSerSerIleAlaAlaGlyValGlnPheGlyTrpAla 47

Qy        383 GCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATG 442
                    ::::::   :::   |||    :::       :::|||:::       |||:::
Db         48 LeuGlnLeuSerLeuLeuThrProTyrValGlnLeuLeuGlyIleProHisLysTrpAla 67

Qy        443 ACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCA 502
              :::::::::          ||||||:::    |||::::::        |||::::::|||
Db         68 SerLeuIleTrpLeuCysGlyProIleSerGlyMetLeuValGlnProIleValGlyTyr 87

Qy        503 GCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCC 562
                    ||||||              |||:::||||||||||||||||||||||     |||      :::
Db         88 HisSerAspArgCysThrSerArgPheGlyArgArgArgProPheIleValAlaGlyAla 107

Qy        563 TTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTG--- 619
                    :::      :::::::::||||||||||              |||:::   |||    :::
Db        108 GlyLeuValThrValAlaValPheLeuIle---------GlyTyrAlaAlaAspIleGly 124

Qy        620 -----------------CTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATC 661
                                    |||         |||:::         :::|||:::
Db        125 HisSerMetGlyAspGlnLeuAspLysProProLysThrArgAlaIleAlaIlePheAla 144

Qy        662 CTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTG 721
              ||||||          :::||||||                           |||      |||
Db        145 LeuGlyPheTrpIleLeuAspValAlaAsnAsnThrLeuGlnGlyProCysArgAlaPhe 164

Qy        722 CTCTCTGACCTCTTCCGG---GACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCC 778
              |||:::||||||             :::              |||   |||    :::   ::::::
Db        165 LeuAlaAspLeuSerAlaGlyAsnAlaLysLysThrArgThrAlaAsnAlaPhePheSer 184

Qy        779 TTCATGATCAGTCTTGGGGGCTGCCTGGGCTAC--------------------CTCCTG 817
              |||    :::::::::|||        |||||||||                        |||
Db        185 PhePheMetAlaValGlyAsnValLeuGlyTyrAlaAlaGlySerTyrArgAsnLeuTyr 204

Qy        818 CCTGCCATTGACTGGGAC---ACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 874
                 :::    :::      |||    :::             |||                  :::
Db        205 LysValValProPheThrMetThrGluSerCysAspLeuTyrCysAlaAsnLeuLysThr 224

Qy        875 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 934
              |||    |||          :::|||||||:::      |||                    :::|||
Db        225 CysPhePheLeuSerIleThrLeuLeu---LeuIleValThrPheValSerLeuCysTyr 243

Qy        935 GCTGAGGAGGCAGCGCTGGGCCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 994
                 :::|||           |||        |||:::|||      :::
Db        244 ValLysGluLysProTrpThrProGluProThrAlaAspGlyLysAla------------ 259

Qy        995 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1054
                                        :::    :::     |||          |||      :::
Db        260 -----------------------SerAsnValProPhe------PheGlyGluIlePhe 269

Qy       1055 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1114
                             :::|||              |||     :::        |||     :::
Db        270 GlyAlaPheLysGluLeu------------LysArgProMetTrpMetLeuLeuIleVal 285

Qy       1115 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 1174
                          :::|||:::|||           |||     ||||||     |||||||:::::::|||
Db        286 ThrAlaLeuAsnTrpIleAlaTrpPheProPheLeuLeuPheAspThrAspTrpMetGly 305

Qy       1175 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 1234
                    :::|||    |||           |||               :::::::::    |||
Db        306 ArgGluValTyrGlyGlyAsnSerAspAlaThrAlaThrAlaAlaSerLysLysLeuTyr 325

Qy       1235 GATGAAGGCGTTCGGATGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTC 1294
                    ::::::|||||||||      |||:::|||||||||     |||              :::
Db        326 AsnAspGlyValArgAlaGlyAlaLeuGlyLeuMetLeuAsnAlaIleValLeuGlyPhe 345

Qy       1295 TTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCC 1354
                 ||||||   ::::::   :::   ::::::   |||      |||      |||
Db        346 MetSerLeuGlyValGluTrpIleGlyArgLysLeuGly---GlyAlaLysArgLeuTrp 364

Qy       1355 AGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTG 1414
                 :::       |||   :::                |||   |||||||:::    :::|||   |||
Db        365 GlyIleValAsnPheIleLeu---------AlaIleCysLeuAlaMetThrValValVal 381

Qy       1414 ------------------------------------------------------------ 1414

Db        382 ThrLysGlnAlaGluAsnHisArgArgAspHisGlyGlyAlaLysThrGlyProProGly 401

Qy       1415 ---GTGACAGCTTCAGCC------------GCCCTCACCGGG------------TTCACC 1447
                 |||||||||   |||                |||:::   |||                   |||
Db        402 AsnValThrAlaGlyAlaLeuThrLeuPheAlaIleLeuGlyIleProGlnAlaIleThr 421

Qy       1448 TTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAG--- 1504
                 ||||||             :::|||:::    |||||||||::::::
Db        422 PheSer------------IleProPheAlaLeuAlaSerIlePheSerThrAsnSerGly 437

Qy       1505 ----------------------------------CAGGTGTTCCTGCCCAAATACCGA 1528
                                                    :::    :::|||:::
Db        438 AlaGlyGlnGlyLeuSerLeuGlyValLeuAsnLeuAlaIleValValProGlnMetVal 457

Qy       1529 GGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCT 1588
                       ||||||         ::::::               :::|||
Db        458 IleSerValGlyGlyGlyProPheAspGluLeuPheGlyGlyGlyAsnIleProAla--- 476

Qy       1589 AAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTC--- 1645
                                   |||    |||    :::    |||      ||||||::::|||
Db        477 ---------------PheValLeuGlyAlaIleAlaAlaAlaValSerGlyValLeuAla 491

Qy       1646 -----------------CCACCTCCACCCGCG 1660
                                 |||||||||   |||
Db        492 LeuThrValLeuProSerProProProAspAla 502
```

RESULT 27
I49607
procollagen type V alpha 2 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 13-Aug-1999
C;Accession: I49607
R;Andrikopoulos, K.; Suzuki, H.R.; Solursh, M.; Ramirez, F.
Dev. Dyn. 195, 113-120, 1992
A;Title: Localization of pro-alpha 2(V) collagen transcripts in the tissues of the de
A;Reference number: I49607; MUID:93214071; PMID:1297453
A;Accession: I49607
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1497 <RES>
A;Cross-references: GB:L02918; NID:g309180; PIDN:AAA37440.1; PID:g309181
C;Genetics:
A;Gene: Col5a-2
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo
F;39-98/Domain: von Willebrand factor type C repeat homology <VWC>
F;1270-1497/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

```
Alignment Scores:
Pred. No.:                5.79e-12     Length:         1497
Score:                    323.00       Matches:        333
Percent Similarity:       30.43%       Conservative:   59
Best Local Similarity:    25.85%       Mismatches:     442
Query Match:              5.19%        Indels:         455
DB:                       2            Gaps:           74

US-09-759-143-110 (1-3410) x I49607 (1-1497)

Qy       3220 TGCAAACGGCACTTAAACCCCCCCTGAGAGATAAGACCTCCCTTAGCTCAGGCAGGGGGT 3161
                 |||       :::   ||||||   |||          |||             |||   ||||||
Db         81 CysAlaAsnProIleThrProProGlyGluCysCysProValCysProGlnThrGlyGly 100
```

```
Qy   3160 GCTCCTGAGTTTCTGTGTGAGATTCCCCAAGCACAGATATACTCTGGGGGCTGAGATGGA 3101
                                                   :::      |||      |||
Db    101 GlyAsp---------------------------ThrSerPheGlyArgGlyArgLysGly 111

Qy   3100 CAAAGGCTTGGGAAACCGCACTTTGTGCTTCTGGTCCTGCAGTAGCTCCAAACAGGGTTG 3041
          |||:::   |||:::|||     |||    :::|||              ||||||:::
Db    112 GlnLys---GlyGluProGlyLeuValProValVal---------------ThrGlyIle 125

Qy   3040 TGGAGCTGGTGGGGAAAGTTGGGG---GTAGGGGAAAGTTGGGGTAGGGGAAATTTTGG 2984
                       |||     |||      |||        |||    ||||||:::   ||
Db    126 ArgGlyArgProGlyProAlaGlyProProGlySerGln-GlyProArgGlyAspArgGl 145

Qy   2983 GCAGTGCCTTCATCAGCCCAGTCCTAGAGAGAGTAGAGGGGAGTGGAAGTGGGGGGAACC 2924
          |            :::|||    ||||||        :::|||          ||||||||
Db    145 yPro---LysGlyArgProGlyProArgGlyProGlnGly------IleAspGlyGluPr 162

Qy   2923 AGGCTGGGCCAAGAGAAGAGGGGTGGTTAGGGAAGCCGTTGAGACCTGAAGCCCCACCCT 2864
          ||||   |||    :::   |||       |||   |||        |||      |||
Db    162 oGlyMet-ProGlyGlnProGlyAlaProGlyProProGlyHisProSerHisProGlyP 182

Qy   2863 CTACCTTCCTTCAACACCCT------------------------------------AACC 2840
             :::      |||                                           |
Db    182 roAspGlyMetSerArgProPheSerAlaGlnMetAlaGlyLeuAspGluLysSerGlyL 202

Qy   2839 TTGGGTAACAGCATTTGGAATTATCATTTGGGATGAGTAGAATTTCCAAGGTCCTGGG-- 2782
          |||||   |||      |||::::::                          ||||||
Db    202 euGlySerGln--ValGlyLeuMet---------------------------ProGlySe 212

Qy   2781 ---------------------------TTAGGCATTTTGGGGGGCCAG------------ 2761
                                     :::|||:::     |||   |||
Db    212 rValGlyProValGlyProArgGlyProValGlyLeuGlnGlyGlnGlnGlyGlyAlaGl 232

Qy   2760 ----------ACCCCAGGAGAAGAA---GATTCTGGCAATGATCAGCCCAATGACCAGCT 2714
                      |||||||||       :::   |||    :::   |||:::   |||
Db    232 yProAlaGlyProProGlyGluProGlyGluProGly-ProMetGlyProIleGlySer- 251

Qy   2713 ATCTCAGGGGACCTGATTGTTGGGGATCCCCCACCCTACCCAAATATTAGACACCAACAC 2654
                |||||||||:::      .|||   |||      |||
Db    252 -----ArgGlyProGlu-----GlyProProGlyLysProGlyGlu-------------- 263

Qy   2653 AGAAAAGCTAGCAATGGATTCCCTTCTACTTTGTTAAATAAATAAGTTAAATATTTAAAT 2594

Db    263 ------------------------------------------------------------ 263

Qy   2593 GCCTGTGTCTCTGTGATGGCAACAGAAGGACCAACAGGCCACATCCTGATAAAAGGTAAG 2534
                        ||||||                    |||    :::
Db    264 --------------AspGly---------------------GluProGlyArgAsnGlyA 272

Qy   2533 AGGGGGGTGGATCAGCAAAAAGACAGTGCTGTGGGCTGAGGGGA---CCTGGTTCTTGTG 2477
             |||                           |||   ||||||   ||||||:::
Db    272 snThrGlyGluValGlyPheSerGlySerProGlyAlaArgGlyPheProGlyAlaProG 292

Qy   2476 TGTTGCCCCTCAGGACTCTTCCCCTACAAATAAGTCATATGTTCAAATCCCATGGAGGAG 2417
           ||||||                                       |||   |||   |||
Db    292 lyLeuPro------------------------------GlyLeuLysGlyHisArgGlyH 302

Qy   2416 TGTTTCATCCTAGAAACTCCCATGCAAGAGCTACATTAAACGAAGCTGCAGGTTAAGGGG 2357
                   ||||||   |||      |||:::
Db    302 isLysGly-LeuGluGlyProLysGlyGluIle--------------------------- 312

Qy   2356 CTTAGAGATGGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGT 2297
                   |||   ||||||            :::   |||      |||
Db    313 ---------GlyAlaProGlyAlaLysGlyGluAlaGlyProThrGlyProMetGlyAla 329

Qy   2296 GTGTCTCAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCAGAGTCC 2237
          :::   ||||||
Db    330 MetGlyProLeuGly--------------------------------------------- 334

Qy   2236 CCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACCCCC 2177
                               |||   |||:::|||            |||
Db    335 ---------------------ProArgGlyMetPro------------GlyGluArgGly 343

Qy   2176 TTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGG------AGAGGGACG 2123
             |||   |||   |||:::|||         |||:::|||         |||
Db    344 ArgLeuGlyProGln---GlyAlaPro---------GlyLysArgGlyAlaHisGlyMet 359

Qy   2122 CCCCAGCCCCCAGCTGTGCAG------CTACGCACCTCAGCAGCACAGGTGGCAGCAGA 2069
          |||   |||         |||                    :::   |||
Db    360 ProGlyLysProGlyProMetGlyProLeuGlyIleProGlySerSerGly--------- 376

Qy   2068 GAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCT 2009
                      |||   |||            :::      |||   ||||||||||||
Db    377 ------------PheProGlyAsnProGlyMetLysGlyGluArgGlyProHisGlyAla 392

Qy   2008 AACAGGAGCGGGGAGCTGGGACCC------------AGTGAGGCAGGCCCTCCACCCCAA 1961
                   |||   ||||||                    |||   |||||||||
Db    393 Arg------GlyProGluGlyProGlnGlyGlnArgGlyGluThrGlyProPro------ 408

Qy   1960 TGTGCTGGAAGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTG 1901
                                        |||            |||   ||||||
Db    409 ------------------------------GlyProAla--GlySerGlnGlyLeuPro- 417

Qy   1900 TGTAGCAAAGTAAATGGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGAC 1841
                                 |||         ||||||   |||:::
Db    418 --------------------GlyAlaValGlyThrAspGlyThrProGlyArgLysGlyA 431

Qy   1840 AGACTGGCTGAGCTGGACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAG 1781
            ||||||   ||||||         |||         ||||||||||||:::      |
Db    431 laThrGlySerAlaGlyThrSerGlyProProGlyLeuAlaGlyProProGlySerProG 451

Qy   1780 GAAGGCAC------------------------------------------------TATC 1769
          ||                                                         |
Db    451 lyProGlnGlySerThrGlyProGlnGlyIleArgGlyGlnSerGlyAspProGlyValP 471

Qy   1768 CAGGATGGCGA---------GGTCCAGGCAGATGCCCCGGCCCGGAA------------- 1731
          ||||::::::            ||||||            |||||||||
Db    471 roGlyPheLysGlyGluAlaGlyProLysGlyGluProGlyProHisGlyIleGlnGlyP 491

Qy   1730 --------CCACCC------------------TGGCCTCGGTGGGCTCACCCACCACCAC 1697
                  ||||||                  ||||||   |||   |||   |||
Db    491 roIleGlyProProGlyGluGluGlyLysArgGlyProArg--GlyAspProGlyThrVa 510

Qy   1696 ACGTACGGAGACATCACAGGCAGAG------GCCCCGCAGAGCGCGGGTGGAGGTGGGAG 1643
                             |||      ||||||   :::   |||      |||||
Db    510 lGlyProProGlyProMetGlyGluArgGlyAlaProGlyAsnArgGlyPheProGlySe 530

Qy   1642 CAGGCCACTGCCT---------------CCAGCACCCACGTGTCCATTAGGGAAGGGAGC 1598
          |       ||||||                                |||:::|||
Db    530 rAspGlyLeuProGlyProLysGlyAlaGlnGlyGluArgGlyProValGlySerSerGl 550

Qy   1597 TCCA------GGCTTAGGG---CCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCT 1547
             |||      |||   |||   |||||||||                  |||   |||
Db    550 yProLysGlyGlyGlnGlyAspProGlyArg-ProGlyGluProGlyLeuProGlyAlaA 570

Qy   1546 AGCACCTCCAGTGTCCCCTCGGTATTTGGGCA---------------GGAACACCTGCTT 1502
                   |||      |||:::                     |||       :
Db    570 rgGlyLeuThrGlyAsnProGlyValGlnGlyProGluGlyLysLeuGlyProLeuGlyA 590

Qy   1501 CTCCCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAA 1442
            ::||||||         |||||||||         ||||||   |||:::|||  |||:::
Db    590 laProGlyGluAspGlyArgPro---GlyProProGlySerIleGlyIleArg-GlyGln 608

Qy   1441 CCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACT---------------GTGGGA 1397
          ||||||      |||                  |||   :::               |||
Db    609 ProGlySerMetGlyValProGlyProLysGlySerSerGlyAspLeuGlyLysProGly 628

Qy   1396 CAGGCATGTGGCACCGGCAG---------------------------------------- 1377
          :::|||         |||
Db    629 GluAlaGlyAsnAlaGlyValProGlyGlnArgGlyAlaProGlyLysAspGlyGluVal 648

Qy   1376 ---CCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTC---GAGTGCCGAATCGCT 1323
             |||   |||            |||   |||      |||   |||         |||
Db    649 GlyProSerGly------ProValGlyProProGlyLeuAlaGlyGluArgGlyGluAla 666

Qy   1322 GCACCAGCCGGT---------------------CCATGACCAGAGAGAAGACCAGGGAGA 1284
```

```
                |||  |||                       |||  |||          ||||||
Db     667 GlyProProGlyProThrGlyPheGlnGlyLeuProGlyProProGlyProProGlyGlu 686

Qy    1283 TGGCGCACTGCAGGAACAGCCCCAGGCTGCCCATCCGAACGCCTTCATCATAGTGTCTCC 1224
                   ||||||       |||  |||  |||
Db     687 GlyGlyLysAlaGlyAsp---GlnGlyValProGlyGlu---------ProGlyAlaVal 702

Qy    1223 GGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTACAGCCCCTCGCCCACGAAAT 1164
           ||||||       |||           ||||||      |||
Db     703 GlyProLeuGlyProArgGlyGluArgGlyAsnProGlyGluArgGlyGluPro------ 720

Qy    1163 CCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGCTGCACAGCTCAGCCACGAAGA 1104
                                                       |||    |||
Db     721 ----------------------------------GlyIleThrGlyLeuProGlyGlu 728

Qy    1103 GCCGGC------GCAGGTGCGGGGCATGCGGC--------------AGCACAGCTGGT 1065
              |||          |||   |||     |||                      |||
Db     729 LysGlyMetAlaGlyGlyHisGlyProAspGlyProLysGlyAsnProGlyProThrGly 748

Qy    1064 GCAGCCGGGGAAGCAGGGCGCCCAGGT---TCCGGAAAGCCAAGC-------------- 1023
              |||:::   |||   ||||||        |||   |||
Db     749 ThrIleGlyAspThrGlyProProGlyLeuGlnGlyMetProGlyGluArgGlyIleAla 768

Qy    1022 ---------GGGCCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTT 972
                     ||||||      |||        |||       ::::|||
Db     769 GlyThrProGlyProLysGlyAspArgGlyGlyIleGlyGluLysGlyAlaGluGlyThr 788

Qy     971 CTG---------CTGGCTCGGTGG------GGCCCAGCGCTGCCTCCTCAGCCACCAGCA 927
                          ||||||         ||||||     ||||||    |||
Db     789 AlaGlyAsnAspGlyAlaArgGlyLeuProGlyProLeuGlyProProGlyProAlaGly 808

Qy     926 GTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGCACT------ 873
           ::                       |||    |||     |||   ||||||
Db     809 LeuLeu---------------GlyAlaProGlyGluProGlyProArgGlyLeuValGly 823

Qy     872 CCTCCTGGGTGCCCAGGT-------------AGGGGG----------------------- 849
           ||||||||     |||                ||||||
Db     824 ProProGlySerArgGlyAsnProGlySerArgGlyGluAsnGlyProThrGlyAlaVal 843

Qy     848 --------------------------------------------------CCAGGGCAC 840
                                                             ||||||
Db     844 GlyPheAlaGlyProGlnGlySerAspGlyGlnProGlyValLysGlyGluProGlyGlu 863

Qy     839 TGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGACTGATCATGA 780
                 |||   |||               ||||||||||||   ||||||
Db     864 ------ProGlyGlnLysGlyAspAlaGlySerProGly---ProGlnGly--------- 877

Qy     779 AGGCATAGACAGAGTAGGCCTGGCGACAGTGGTCCGGGT--------------------- 741
                             ||||||   ||||||
Db     878 ------LeuAlaGlySerProGlyProHisGlyProHisGlyValProGlyLeuLysGly 895

Qy     740 -------------------------------------CCCGGAAGAGGTCAGAGAGCA 720
                                                 ||||||        |||
Db     896 GlyArgGlyThrGlnGlyProProGlyAlaThrGlyPheProGlySerAlaGlyArgVal 915

Qy     719 GGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGCCCAGGA 660
           |||||||||           ||||||       ||||||       ||||||
Db     916 GlyProProGlyProAlaGlyAlaProGlyProAlaGlyProAlaGly---GluProGly 934

Qy     659 TGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGG 600
              |||                 |||||||||:::           :::
Db     935 LysGluGlyProProGlyLeuArgGlyAspProGlySerHisGlyArgValGly------ 952

Qy     599 CCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCAAGGACAGTGCCCAGATGAAGGGCC 540
                       :::||||||   ||||||   |||    :::
Db     953 -----------AspArgGlyProAlaGlyProProGlySer------------------ 962

Qy     539 GGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTAGGAGCGGGACACAGA 480
                                                       |||   :::
Db     963 ---------------------------------------------ProGlyAspLys 966

Qy     479 CCAGGCCCAGCACTGGACCAATGCCCAGCACCATGGTCATGAACTTCTCCTCTACCCCCA 420
              |||     |||     |||   |||                    ||||||
Db     967 GlyAspProGlyGluAspGlyGlnProGlyProAspGly--------------ProPro 981

Qy     419 CTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCCAAACACACCTCCAGGCCAA 360
              ||||||                |||     :::                |||||||||
Db     982 GlyProAlaGlyThrThrGlyGlnArgGlyIleVal---------GlyMetProGlyGln 998

Qy     359 AGG---------------------------TTAGCAGGTTGACCAGCAAGAGCTGGG 330
           |||                                 ||||||   |||   :::   |||
Db     999 ArgGlyValThrGlyMetProGlyLeuProGlyProAlaGlyThrProGlyLysValGly 1018

Qy     329 CTTTCCGGTGCCGCAGCAGGC--------------------GGCTCACCCACAGCCTCT 291
              :::||||||                                 |||   |||   ::::::
Db    1019 ProThrGlyAlaThrGlyAspLysGlyProProGlyProValGlyProProGlySerAsn 1038

Qy     290 GGACCATAGTGGG---CCAGGCGGGTAGGGCTCAGGGGGCCGTTCAGGCACTCCAGAACT 234
           ||||||   |||   ||||||    |||      |||      |||||||||
Db    1039 GlyPro-ValGlyGluProGlyProGluGlyProAlaGlyAsnAspGlyThrProGly-- 1057

Qy     233 GCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAACTGCC 174
                 |||   |||:::      |||
Db    1058 ----ArgAspGlyAlaValGlyGluArgGly-------------------------As 1067

Qy     173 TAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGCTTCTCAGCC 114
              |||      ||||||                    |||   |||      :::
Db    1067 pArgGlyAspProGly--------------ProAlaGlyLeuProGlySerGln-GlyA 1082

Qy     113 CATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTCATCACTCAGATC 54
                     ||||||     ||||||          ||||||               :::|
Db    1082 laProGlyThrProGlyProValGly---AlaProGlyAspAlaGlyGlnArgGlyGluP 1101

Qy      53 CTGGC---CGAGGCGCGCGGCTGTCACCCGGAGCCAGCGCGTGCAGGCTGGTTCCC 1
              |||||   ||||||          ||||||     :::      |||    :::|||
Db    1101 roGlySerArgGlyProValGlyProProGlyArgAlaGlyLysArgGlyLeuPro 1119
```

RESULT 28
S22917
collagen alpha 5(IV) chain precursor, renal splice form - human
N;Alternate names: procollagen alpha 5(IV) chain
N;Contains: collagen alpha 5(IV) chain precursor, leukocyte splice form
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 27-Feb-1997 #text_change 21-Jul-2000
C;Accession: S22917; A54365; A57079; A37122; I54317; A34850; S18850; I56971; I76598;
R;Zhou, J.; Hertz, J.M.; Leinonen, A.; Tryggvason, K.
J. Biol. Chem. 267, 12475-12481, 1992
A;Title: Complete amino acid sequence of the human alpha-5(IV) collagen chain and ide
n Alport syndrome patient.
A;Reference number: S22917; MUID:92316923; PMID:1352287
A;Accession: S22917
A;Molecule type: mRNA
A;Residues: 1-967 <ZHO>
A;Cross-references: GB:M90464; NID:g180826; PIDN:AAA52046.1; PID:g553234
R;Zhou, J.; Leinonen, A.; Tryggvason, K.
J. Biol. Chem. 269, 6608-6614, 1994
A;Title: Structure of the human type IV collagen COL4A5 gene.
A;Reference number: A54365; MUID:94165049; PMID:8120014
A;Accession: A54365
A;Molecule type: DNA
A;Residues: 1-922 <ZH2>
A;Cross-references: GB:U04470; NID:g463378; GB:U04520; NID:g463428; PIDN:AAC27816.1;
R;Zhou, J.; Mochizuki, T.; Smeets, H.; Antignac, C.; Laurila, P.; de Paepe, A.; Trygg
Science 261, 1167-1169, 1993
A;Title: Deletion of the paired alpha5(IV) and alpha6(IV) collagen genes in inherited
A;Reference number: A57079; MUID:93361972; PMID:8356449
A;Accession: A57079
A;Molecule type: DNA
A;Residues: 1-27 <ZH4>
A;Cross-references: GB:Z37153; NID:g587203; PIDN:CAA85512.1; PID:g587204
R;Pihlajaniemi, T.; Pohjolainen, E.R.; Myers, J.C.
J. Biol. Chem. 265, 13758-13766, 1990
A;Title: Complete primary structure of the triple-helical region and the carboxyl-ter
A;Reference number: A37122; MUID:90337990; PMID:2380186
A;Accession: A37122
A;Molecule type: mRNA

A;Residues: 84-439,'GS',442-624,'LALQ',629-666,'FR',669-887,'R',889-1264,1271-1691 <PIH>
A;Cross-references: GB:J05558; EMBL:M58526; NID:g1314209
A;Note: submitted to the EMBL Data Library, February 1991
A;Note: the authors translated the codon GCC for residue 115 as Val
R;Renieri, A.; Seri, M.; Myers, J.C.; Pihlajaniemi, T.; Massella, L.; Rizzoni, G.; De Ma
Hum. Mol. Genet. 1, 127-129, 1992
A;Title: De novo mutation in the COL4A5 gene converting glycine 325 to glutamic acid in
A;Reference number: I54317; MUID:93244772; PMID:1363780
A;Accession: I54317
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 313-324,'E',326-330 <REN>
A;Cross-references: GB:S59334; NID:g299946; PIDN:AAD13909.1; PID:g4261609
R;Hostikka, S.L.; Eddy, R.L.; Byers, M.G.; Hoeyhtyae, M.; Shows, T.B.; Tryggvason, K.
Proc. Natl. Acad. Sci. U.S.A. 87, 1606-1610, 1990
A;Title: Identification of a distinct type IV collagen alpha chain with restricted kidne
A;Reference number: A34850; MUID:90160375; PMID:1689491
A;Accession: A34850
A;Molecule type: mRNA
A;Residues: 914-1264,1271-1691 <HOS>
A;Cross-references: EMBL:M31115; NID:g180824; PIDN:AAA52045.1; PID:g180825
R;Zhou, J.; Hostikka, S.L.; Chow, L.T.; Tryggvason, K.
Genomics 9, 1-9, 1991
A;Title: Characterization of the 3' half of the human type IV collagen alpha-5 gene that
A;Reference number: A37969; MUID:91169491; PMID:2004755
A;Accession: S18850
A;Molecule type: DNA
A;Residues: 924-1264,1271-1691 <ZH3>
A;Cross-references: EMBL:M63455; EMBL:M63456; EMBL:M63457; EMBL:M63458; EMBL:M63459; EMB
8; EMBL:M63470; EMBL:M63471; EMBL:M63472; EMBL:M63473; NID:g177922; PIDN:AAA51558.1; PID
R;Guo, C.; Van Damme, B.; Van Damme-Lombaerts, R.; Van den Berghe, H.; Cassiman, J.J.; M
Kidney Int. 44, 1316-1321, 1993
A;Title: Differential splicing of COL4A5 mRNA in kidney and white blood cells: a complex
A;Reference number: I56971; MUID:94133540; PMID:8301933
A;Accession: I56971
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1258-1276 <GUO1>
A;Cross-references: GB:S69168; NID:g545095; PIDN:AAC60612.1; PID:g545096
A;Note: kidney splice form
A;Accession: I76598
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1284-1291,'TFLGYLACLV' <GUO2>
A;Cross-references: GB:S69169; NID:g545097; PIDN:AAC60613.1; PID:g545098
A;Note: frameshift mutation in patient with Alport syndrome
R;Myers, J.C.; Jones, T.A.; Pohjolainen, E.R.; Kadri, A.S.; Goddard, A.D.; Sheer, D.; Sc
Am. J. Hum. Genet. 46, 1024-1033, 1990
A;Title: Molecular cloning of alpha5(IV) collagen and assignment of the gene to the regi
A;Reference number: A35335; MUID:90252791; PMID:2339699
A;Accession: A35335
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1448-1477 <MYE>
R;Nakazato, H.; Hattori, S.; Ushijima, T.; Matsuura, T.; Koitabashi, Y.; Takada, T.; Yos
Kidney Int. 46, 1307-1314, 1994
A;Title: Mutations in the COL4A5 gene in Alport syndrome: a possible mutation in primord
A;Reference number: I56975; MUID:95156893; PMID:7853788
A;Accession: I56975
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1595-1602 <NAK>
A;Cross-references: GB:S75903; NID:g913882; PIDN:AAB33374.1; PID:g913883
A;Note: permature termination mutation from a patient with Alport syndrome; one other mu
R;Lemmink, H.H.; Schroeder, C.H.; Brunner, H.G.; Nelen, M.R.; Zhou, J.; Tryggvason, K.;
Genomics 17, 485-489, 1993
A;Title: Identification of four novel mutations in the COL4A5 gene of patients with Alpo
A;Reference number: I54188; MUID:94010948; PMID:8406498
A;Accession: I54188
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1604-1607,'VHDAYKC' <LEM>
A;Cross-references: GB:S65767; NID:g425563; PIDN:AAD13967.1; PID:g4261667

A;Note: frameshift mutation from a patient with Alport syndrome; five other mutations
C;Comment: Prolines and lysines at the third position of the tripeptide repeating uni
ed and subsequently O-glycosylated.
C;Genetics:
A;Gene: GDB:COL4A5; ATS
A;Cross-references: GDB:120596; OMIM:303630
A;Map position: Xq22-Xq22
A;Introns: 27/3; 47/3; 77/3; 92/3; 107/3; 128/3; 146/3; 155/3; 182/3; 203/3; 215/3; 2
/3; 799/1; 837/1; 893/1; 923/1; 973/1; 1006/1; 1036/1; 1082/3; 1125/1; 1152/1; 1185/1
A;Note: the alpha 5(IV) and alpha 6(IV) chain genes are encoded on opposite strands w
C;Complex: This minor type IV collagen is thought to form a heterotrimer of two alpha
mong trimer amino-terminal domains (with disulfide and desmosine cross-links), dimeri
er associations in the interrupted helical domain (with disulfide and desmosine cross
C;Function:
A;Description: minor structural component of extracellular basement membrane
C;Superfamily: collagen alpha 1(IV) chain
C;Keywords: Alport syndrome; basement membrane; coiled coil; extracellular matrix; gl
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-1691/Product: collagen alpha 5(IV) chain, renal splice form #status predicted <M
F;27-1264,1271-1691/Product: collagen alpha 5(IV) chain, leukocyte splice form #statu
F;27-41/Domain: amino-terminal nonhelical, NC2 #status predicted <NC2>
F;42-1462/Region: interrupted helical
F;1463-1691/Domain: carboxyl-terminal nonhelical, NC1 #status predicted <NC1>
F;1473-1573/Domain: collagen IV carboxyl-terminal repeat <CT1>
F;1583-1687/Domain: collagen IV carboxyl-terminal repeat <CT2>
F;29,32,38,40,124,451,481,484/Disulfide bonds: interchain #status predicted
F;125/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;1482-1570,1515-1573/Disulfide bonds: (or 1482-1573, 1515-1570) #status predicted
F;1527-1533,1638-1644/Disulfide bonds: #status predicted
F;1592-1684,1626-1687/Disulfide bonds: (or 1592-1687, 1626-1684) #status predicted

```
Alignment Scores:
Pred. No.:              5.79e-12      Length:        1691
Score:                  323.00        Matches:       284
Percent Similarity:     30.90%        Conservative:  59
Best Local Similarity:  25.59%        Mismatches:    339
Query Match:            5.03%         Indels:        428
DB:                     1             Gaps:          65

US-09-759-143-110 (1-3410) x S22917 (1-1691)

Qy          2 GGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCT-----------CGG 49
                ||||||   :::           ||||||          |||
Db        486 GlyThrGlyIleSerGly---ProProGlyGlnProGlyLeuProGlyLeuProGlyPro 504

Qy         50 CCAGGATCTGAGTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAG 109
              |||||||||                    |||     |||         |||   |||
Db        505 ProGlySer--------------------LeuGlyPheProGlyGlnLysGly---Glu 516

Qy        110 CATGGGCTGAGAAGCTGGACCGGCACCAAAGGGCTG--------------GCAGAAATG 154
                 |||         ||||||  |||||||||                 |||
Db        517 LysGlyGlnAlaGlyAlaThrGlyProLysGlyLeuProGlyIleProGlyAla-ProGl 536

Qy        155 GGCGCCTGGCTGATTCCTAGGCAGTTGGCGGCAGCAAGGA------------------- 194
                |||||||||   |||   ||||||       :::    |||
Db        536 yAlaProGly---PheProGlySerLysGlyGluProGlyAspIleLeuThrPheProGl 555

Qy        195 ----------------GGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTC 238
                                ||||||      :::    ||||||
Db       '555 yMetLysGlyAspLysGlyGluLeuGlySerProGlyAlaPro---------GlyLeuPr 572

Qy        239 TGGAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTG 298
                |||:::|||       |||     :::            ||||||      |||||||:::
Db        572 oGlyLeuProGlyThrProGlyGlnAspGlyLeuProGly-LeuProGlyProLys---- 590

Qy        299 TGGGTGAGCCG--------------------CCTGCTGCGGCACCGGAAAGCCCAGCTC 337
                |||||||||                                  |||    :::|||   |
Db        591 --GlyGluProGlyGlyIleThrPheLysGlyGluArgGlyProProGlyAsnProGlyL 610

Qy        338 TTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTAT 397
               ||    |||    |||    |||:::    |||    |||
Db        610 euProGlyLeuProGlyAsnIleGlyProMetGly------------------------ 621
```

```
Qy      398 GTGCCGCCTCTGCTGCTGGAAGTGGGGTAGAGGAGAAGTTCATGACCATGGT------- 450
                        ||| |||      |||   |||  :::  |||
Db      622 -----------ProProGlyPheGlyProProGlyProValGlyGluLysGlyIleGlnG 638

Qy      451 -----GCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCC 505
                 ||||||      :::      |||||||:::   |||   |||:::
Db      638 lyValAlaGly------AsnProGlyGlnProGlyIle---ProGlyProLys------- 652

Qy      506 AGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTG 565
                                                                       |
Db      653 -----------------------------------------------------------G 653

Qy      566 GGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGC 625
             ||   |||    :::   ::::::   |||                  :::     :::|
Db      653 lyAspProGlyGlnThrIleThrGlnPro---------------GlyLysProGlyLeuP 668

Qy      626 CCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGG-----------CGTGGGG 673
            ||||||:::                                    ||||||       |||
Db      668 roGlyAsnProGlyArgAspGlyAspValGlyLeuProGlyAspProGlyLeuProGlyG 688

Qy      674 CTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCT- 732
                  ||||||      ||||||:::      |||   |||   |||   :::   |||
Db      688 lnProGlyLeu------ProGlyIleProGlySerLysGlyGluProGlyIle---ProG 705

Qy      733 --------CTTCCGGGA---CCCGGACCACTGTCG---CCAGGCCTACTCTGTCTATGCC 778
                    |||||||||   |||||||||         ||||||:::
Db      705 lyIleGlyLeuProGlyProProGlyProLysGlyPheProGlyIlePro---------- 721

Qy      779 TTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACC 838
                                |||   ||||||   |||         :::|||   :
Db      722 ----------------GlyProProGlyAlaProGlyThrProGlyArgIleGlyLeuG 736

Qy      839 AGTGCCCTGGCCCCCTACCTGGGCACCCAGGA--------------------GGAGTGC 877
            ::   |||      ||||||   ||||||                        :::|
Db      736 luGlyProProGlyProProGlyPheProGlyProLysGlyGluProGlyPheAlaLeuP 756

Qy      878 CTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCT 937
             ||      |||      |||   ||||||   :::   :::               |||
Db      756 roGlyProProGlyProProGlyLeuProGlyPheLysGlyAlaLeuGlyProLysGlyA 776

Qy      938 GAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCG 997
                  |||      ||||||                           ||||||
Db      776 spArgGlyPheProGlyPro--------------------ProGlyProProGlyArgT 789

Qy      998 CCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCC 1057
                |||      :::|||||||||   |||         |||                |
Db      789 hrGlyLeuAspGlyLeuProGlyProLysGlyAspValGlyProAsnGlyGlnProGlyP 809

Qy     1058 CGGCTGCACCAGCTGTGCTGCCG-----------CATGCCCCGCACCCTGCGCCGGCTC 1105
             ||      |||      ||||||               |||         |||    :
Db      809 roMetGlyProProGlyLeuProGlyIleGlyValGlnGlyProProGlyProProGlyI 829

Qy     1106 TTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGAT 1165
            ::   |||      |||                           |||||||||
Db      829 leProGlyProIleGlyGln--------------------------ProGlyLeuHisGly- 840

Qy     1166 TTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGG 1225
                             :::||||||               |||      ||||||
Db      841 -----------------IleProGlyGlu------------LysGlyAspProGlyProP 851

Qy     1226 AGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATC 1285
               |||                         ||||||   |||         |||
Db      851 roGlyLeu-----------------AspValProGlyPro---ProGlyGluArgGlyS 864

Qy     1286 TCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTC 1345
               ||||||:::            ||||||   |||               |||    |
Db      864 erProGlyIleProGlyAlaProGlyProIleGlyProPro------GlySerProGlyL 882

Qy     1346 TATTTGGCCAGTGTGGCAGC------TTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCC 1399
             ||   |||:::   |||:::          ||||||   |||   :::   |||   :::
Db      882 euProGlyLysAlaGlyAlaSerGlyPhePro---GlyThrLysGlyGluMetGlyMet- 900

Qy     1400 CACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTG 1459

Db      900 ------------------------------------------------------------ 900

Qy     1460 CAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCC 1519
                                    |||   |||   |||||||||   |||:::|||   :
Db      901 ----------------MetGlyProProGlyProProGlyProLeuGlyIleProGlyA 915

Qy     1520 AAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTG 1579
            ::       |||      |||         :::|||   |||
Db      915 rg------SerGly-ValProGlyLeuLysGlyAspAspGlyLeuGln----------- 928

Qy     1580 CCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGC 1639
                  :::||||||   |||   |||   |||      |||:::   |||
Db      929 ------GlyGlnProGlyLeuProGlyProThrGlyGluLysGlySerLysGlyGluPro 946

Qy     1640 CTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTG 1699
              ||||||   |||  |||            ||||||                |||   |||   :::
Db      947 GlyLeuProGlyPro-Pro---------GlyPro---------MetAspProAsnLeuLe 960

Qy     1700 GTGGTGGGTGAGCCCACCGAGGCCAGGGTGGT--------------------TCCGGGC 1738
              |||         :::|||:::   |||                          ||||||
Db      960 u-GlySerLysGlyGluLysGlyGluProGlyLeuProGlyIleProGlyValSerGlyP 980

Qy     1739 CG---GGGCAT------CTGCCTGGA---CCTCGCCATCCTGGATAGTGC---CTTCCTG 1783
             ||   |||:::      |||||||||   |||         ||||||         |||
Db      980 roLysGlyTyrGlnGlyLeuProGlyAspProGlyGlnProGlyLeuSerGlyGlnProG 1000

Qy     1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
               :::||||||   |||   |||                  |||   ||||||
Db     1000 lyLeuProGlyProProGlyProLysGlyAsnProGlyLeuProGlyGlnPro------- 1017

Qy     1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
                        |||:::         ||||||                          |
Db     1018 -----------GlyLeuIleGlyProProGly--------------------LeuLysG 1027

Qy     1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTG 1963
             ||:::                                                  ||||
Db     1027 lyThr----------------------------------------------------IleG 1030

Qy     1964 GGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTG 2023
             ||      |||   |||         :::   |||      |||   |||:::
Db     1030 lyAspMetGlyPhePro-GlyProGlnGlyValGluGlyProProGlyProSerGlyVal 1049

Qy     2024 CCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTG 2083
            ||||||   |||         ||||||
Db     1050 ProGlyGlnProGlySerProGly-LeuProGly------------------------ 1060

Qy     2084 TGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCA 2143
                                      |||      |||   |||   :::|||
Db     1061 ---------------------GlnLysGlyAspLysGlyAspProGlyIleSer----- 1071

Qy     2144 GTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGG---GTTTCAGTCTGGACTTATAC 2200
              |||:::|||||||||                 ||||||         :::      |||
Db     1072 -SerIleGlyLeuProGlyLeuProGlyProLysGlyGluProGlyLeuProGlyTyrPr 1091

Qy     2201 AGGGAGGCCA----------------------GAAGGGCTCCATGCACTGGAATGCGG 2236
              |||   |||                                ||||||      |||
Db     1091 oGlyAsnProGlyIleLysGlySerValGlyAspProGlyLeuProGlyLeuProGlyTh 1111

Qy     2237 GGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACAC 2296
                  |||   |||   |||   |||
Db     1111 rProGlyAlaLysGlyGlnProGlyLeu------------------------------- 1120

Qy     2297 ACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAG 2356
                     ||||||   |||                   :::||||||   |||
Db     1121 -------ProGlyPheProGly--------------ThrProGlyProProGlyProLy 1133

Qy     2357 CCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACA 2416
                                                      |||:::|||   |||
Db     1133 s-------------------------------------GlyIleSer-Gly------- 1137

Qy     2417 CTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACA 2476
             ||||||   :::                                          |||
```

```
Db       1138 --ProProGlyAsn-----------------------------------ProGlyLeuP 1145

Qy       2477 CACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTT 2536
                   |||||||||||:::    .                             ||||||
Db       1145 roGlyGluProGlyProValGlyGlyGlyGlyHisProGlyGlnProGlyProPro---- 1163

Qy       2537 ACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATT 2596

Db       1163 ------------------------------------------------------------ 1163

Qy       2597 TAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTG 2656

Db       1163 ------------------------------------------------------------ 1163

Qy       2657 TTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGC 2716
                                 |||      |||    |||    |||   |||
Db       1164 -----------------GlyGluLysGlyLysProGlyGlnAspGly------------ 1173

Qy       2717 TGGTCATTGGGCTGATCATTGCCA------------GAATCTTCTTCTCCTGGGGTCTGG 2764
                                 ||| |||            :::       ||||||
Db       1174 -----------------Ile-ProGlyProAlaGlyGlnLysGlyGluProGly------ 1185

Qy       2765 CCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCA 2824
                 |||         |||||||||||                             |||
Db       1186 GlnProGlyPheGlyAsnProGlyPro-----------------------------Pro 1195

Qy       2825 AATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGT------GGGGCTTCAG 2878
                                   |||:::   |||  :::   |||      |||      |
Db       1196 --------------------GlyLeuProGly--LeuSerGlyGlnLysGlyAspGlyG 1208

Qy       2879 GTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTC 2938
               |||||   |||    |||     :::|||          ||||||    |||   |||  |||
Db       1208 lyLeuProGlyIleProGlyAsnProGlyLeuProGlyPro-LysGlyGluProGlyPhe 1227

Qy       2939 CAC------------------TCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCAC 2980
               |||                   |||
Db       1228 HisGlyPheProGlyValGlnGlyPro------------------------------- 1236

Qy       2981 TGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCA 3040
                          |||  ||||||     |||   |||   |||      |||
Db       1237 ------------ProGlyProProGlySerProGlyProAlaLeuGluGlyProLysGly 1252

Qy       3041 CAACCC 3046
                 |||
Db       1253 AsnPro 1254
```

RESULT 29
CGBO1S
collagen alpha 1(I) chain - bovine (tentative sequence) (fragments)
C;Species: Bos primigenius taurus (cattle)
C;Date: 24-Apr-1984 #sequence_revision 31-Dec-1993 #text_change 31-Mar-2000
C;Accession: A91193; A91229; A91387; A91211; A91201; A91200; A43048; A02853
R;Rauterberg, J.; Timpl, R.; Furthmayr, H.
Eur. J. Biochem. 27, 231-237, 1972
A;Title: Structural characterization of N-terminal antigenic determinants in calf and hu
A;Reference number: A91193; MUID:72255334; PMID:4115172
A;Accession: A91193
A;Molecule type: protein
A;Residues: 1-19 <RAU>
A;Experimental source: skin
A;Note: the epsilon carbon of Lys-9, by homology with the rat alpha 1(I) chain, is conve
R;Fietzek, P.P.; Kuehn, K.
Eur. J. Biochem. 52, 77-82, 1975
A;Title: The covalent structure of collagen: amino-acid sequence of the cyanogen-bromide
A;Reference number: A91229; MUID:76022320; PMID:1164916
A;Accession: A91229
A;Molecule type: protein
A;Residues: 20-145 <FIE>
A;Experimental source: skin
A;Note: Lys-103 is hydroxylated and binds glucosylgalactose
R;Fietzek, P.P.; Wendt, P.; Kell, I.; Kuehn, K.
FEBS Lett. 26, 74-76, 1972
A;Title: The covalent structure of collagen: amino acid sequence of alpha1-CB3 from calf
A;Reference number: A91387; MUID:73049499; PMID:4673951
A;Accession: A91387
A;Molecule type: protein
A;Residues: 146-294 <FI2>
A;Experimental source: skin
R;Fietzek, P.P.; Rexrodt, F.W.; Hopper, K.E.; Kuehn, K.
Eur. J. Biochem. 38, 396-400, 1973
A;Title: The covalent structure of collagen. 2. The amino-acid sequence of alpha1-CB7
A;Reference number: A91211; MUID:74086118; PMID:4359390
A;Accession: A91211
A;Molecule type: protein
A;Residues: 295-562 <FI3>
A;Experimental source: skin
R;Wendt, P.; Mark, K.V.D.; Rexrodt, F.; Kuehn, K.
Eur. J. Biochem. 30, 169-183, 1972
A;Title: The covalent structure of collagen. The amino-acid sequence of the 112 resid
A;Reference number: A91201; MUID:73042276; PMID:4343808
A;Accession: A91201
A;Molecule type: protein
A;Residues: 563-675 <WEN>
A;Experimental source: skin
R;Fietzek, P.P.; Rexrodt, F.W.; Wendt, P.; Stark, M.; Kuehn, K.
Eur. J. Biochem. 30, 163-168, 1972
A;Title: The covalent structure of collagen. Amino acid sequence of peptide alpha1-CB
A;Reference number: A91200; MUID:73042275; PMID:4343807
A;Accession: A91200
A;Molecule type: protein
A;Residues: 676-758 <FI4>
A;Experimental source: skin
A;Note: Pro-726 is the only 3-hydroxyproline and the only hydroxylated proline in pos
R;Rauterberg, J.; Fietzek, P.; Rexrodt, F.; Becker, U.; Stark, M.; Kuehn, K.
FEBS Lett. 21, 75-79, 1972
A;Title: The amino acid sequence of the carboxyterminal nonhelical cross link region
A;Reference number: A43048
A;Accession: A43048
A;Molecule type: protein
A;Residues: 759-779 <RA2>
A;Experimental source: skin
C;Comment: Lysines at positions 115, 124, 274, 346, 424, 496, 658, and 670 may be hyd
C;Comment: Prolines in the third position of the tripeptide repeating unit (G-X-Y) ar
C;Comment: The order of the eight CNBr peptides in the alpha 1(I) chain of bovine ski
9, 149, 268, and 217 residues.
C;Comment: The complete chain contains 1052 residues.
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo
C;Keywords: coiled coil; extracellular matrix; glycoprotein; pyroglutamic acid; trime
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

```
Alignment Scores:
Pred. No.:                7.65e-12      Length:         779
Score:                    321.00        Matches:        229
Percent Similarity:       31.12%        Conservative:   38
Best Local Similarity:    26.69%        Mismatches:     269
Query Match:              5.16%         Indels:         324
DB:                       1             Gaps:           51
```

US-09-759-143-110 (1-3410) x CGBO1S (1-779)

```
Qy       2365 GTTAAGGGGCTTAGA---------------GATGGGAAACCAGGTGACTGAGTTTATTCA 2321
              :::||||||   |||                    ||||||     ||||||        :::
Db        102 MetLysGlyHisArgGlyPheSerGlyLeuAspGlyAlaLysGlyAsp---------Ala 118

Qy       2320 GCTCCCAAAAACCCTTCTCTAGGTGTGTCTCAACTAGGAGGCTAGCTGTTAACCCTGAGC 2261
                 |||     |||
Db        119 GlyProAlaGlyProLys------------------------------------------ 124

Qy       2260 CTGGGTAATCCA------CCTGCAGAGTCCCCGCATTCCAGTGCATGGAGCCCTTCTGGC 2207
                 |||   |||      |||   |||:::                      |||   |||
Db        125 ---GlyGluProGlySerProGlyGluAsnGlyAlaProGlyGlnMetGlyProArgGly 143

Qy       2206 CTCCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGA 2147
              ||||||                  |||                      :::|||   |||
Db        144 LeuPro---------------GlyPheProGlyProLysGlyAlaAlaGlyGluProGly 158
```

```
Qy     2146 GACTGGGGAGAGAGGAGAGGGACGCCCCAGCCCCCA---GCTGTGCAGCTACGCACCTCA 2090
                 ||||||   ||||||   |||   ||||||   ||||||
Db      159 LysAlaGlyGlu---ArgGlyValProGlyProProGlyAlaValGlyProAlaGlyLys 177

Qy     2089 GCAGCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCA 2030
               :::    |||::::::  |||
Db      178 AspGlyGluAlaGlyAlaGlnGlyPro--------------------------------- 186

Qy     2029 GCCCGGCAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGGA------------------ 1988
                      |||    |||     :::||||||  |||
Db      187 ---------ProGlyProAlaGlyProAlaGlyGluArgGlyGluGlnGlyProAlaGly 203

Qy     1987 ---------------------CCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGT 1949
                                 |||     |||||||||||    :::    ||||||
Db      204 SerProGlyPheGlnGlyLeuProGlyProAlaGlyProProGlyGlu---AlaGlyLys 222

Qy     1948 TTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTA 1889

Db      222 ------------------------------------------------------------ 222

Qy     1888 AATGGCGACCAGAC----CCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGC 1833
              |||:::|||    ||||||          |||
Db      223 ProGlyGluGlnGlyValProGlyAspLeuGlyAlaPro--------------------- 235

Qy     1832 TGAGCTGGACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGA---AGG 1776
                                             ||||||   |||    |||   |||
Db      236 ---------------------------------GlyProSerGlyAlaArgGlyGluArg 244

Qy     1775 CACTATCCAGGATGGCGAGGTCCAGGCAGATGCCCCGGCCCGGAACCACCCTG-GCCTCG 1717
              :::||||||   ||||||          |||||||||    |||      :::
Db      245 GlyPheProGlyGluArgGlyValGluGlyProProGlyProAlaGlyProArgGlyAla 264

Qy     1716 GTGGGCTCACCCACCACC------ACACGTACGGAGACATCACAGGCAGAGGCCCCGCAG 1663
              |||:::|||           :::   :::            ||||||
Db      265 AsnGlyAlaProGlyAsnAspGlyAlaLysGlyAspAlaGlyAlaProGlyAlaProGly 284

Qy     1662 AGCGCGGGTGGAGGTGGGAGCAGGCCACTGCCT------------------------CCA 1627
             |||   |||    |||   :::   :::|||                         |||
Db      285 SerGlnGlyAlaProGlyLeuGlnGlyMetProGlyGluArgGlyAlaAlaGlyLeuPro 304

Qy     1626 GCACCCACGTGTCCATTAGGG------------AAGGGAGCTCCAGGCTTAGGGCCTGGC 1579
              |||           |||        |||||||||  |||    ||||||
Db      305 GlyProLysGlyAspArgGlyAspAlaGlyProLysGlyAlaAspGly---AlaProGly 323

Qy     1578 AGGAAGCTGGTCATCAGGCTGTCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGTATT-- 1521
             :::      |||    |||:::       :::  ||||||   |||   |||
Db      324 LysAspGlyValArgGlyLeuThrGlyProIleGlyProPro-GlyProAlaGlyAlaPr 343

Qy     1520 ------------------------------TGGGCAGGAACACCTGCTTCTCCCGGTGG 1492
                                          |||||||||      :::||||||
Db      343 oGlyAspLysGlyGluAlaGlyProSerGlyProAlaGlyThrArgGlyAlaProGlyAs 363

Qy     1491 TAGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGG 1432
              ||||||   |||             |||    |||
Db      363 pArgGlyGluPro------------GlyProProGly----------------------- 371

Qy     1431 GCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACCGGCAGCCACA 1372
                                                                   |||
Db      372 -------------------------------------------------------ProAl 373

Qy     1371 GGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCTGCACCAGCCGG 1312
             |||                                          |||   |||   ||
Db      373 aGly-----------------------------------------PheAlaGlyProProGl 380

Qy     1311 TCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACAGCCCCAGGCTGCCC 1252
             |         :::   :::||||||                    |||   |||   ||
Db      380 y------AlaAspGlyGlnProGly--------------------AlaLysGlyGluPr 391

Qy     1251 ATCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACG 1192
             |    :::                :::||||||       ||||||
Db      391 oGlyAspAlaGlyAlaLysGlyAspAlaGlyProProGlyProAlaGlyProAlaGlyPr 411

Qy     1191 CCCTGGTACA----------------GCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTG 1147
             ||||||                         ||||||   |||            |||   |||
Db      411 oProGlyProIleGlyAsnValGlyAlaProGlyPro-----------LysGlyAlaAr 427

Qy     1146 AAGGTCATGAGTGCCATCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGG 1087
               |||      |||           |||       |||                    ||
Db      427 gGlySerAlaGlyProProGlyAlaThrGlyPhePro--------------------Gl 440

Qy     1086 GGCATGCGGCAGCACAGCT------GGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCCGG 1033
             ||||    |||                |||   :::|||   ||||||   ||||||
Db      440 yAlaAlaGlyArgValGlyProProGlyProSerGlyAsnAlaGlyProProGly----- 458

Qy     1032 AAAGCCAAGCGGGCCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCT 973
                |||    ||||||                            ||||||    |||
Db      459 ----ProProGlyProAlaGlyLysGluGlySerLysGlyProArgGlyGluThrGly-- 476

Qy      972 TCTGCTGGCTCGGTGGGGCCCAGCGCTGCCTCCTCAGCCACCAGCAGTGTGGCTGCTACG 913
                     |||    |||            ||||||   ||||||
Db      477 -ProAlaGlyArgProGlyGluValGlyProProGlyProPro----------------- 490

Qy      912 CAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGCACTCCT----------------- 870
                                 |||   |||||:::::|||   |||
Db      491 ----------------GlyProAlaGlyGluLysGlyAlaProGlyAlaAspGlyProAl 505

Qy      869 -------CCTGGGTGCCCAGGT--------------------AGGGGG------------ 849
                    ||||||   ||||||                    ||||||
Db      505 aGlyAlaProGlyThrProGlyProGlnGlyIleAlaGlyGlnArgGlyValValGlyLe 525

Qy      848 -CCAGGGCACTGGTGTCCCAGTCAATGGCAGGAGGA------GGTAGCCCAGGCAGCCC 796
             ||||||                    |||   |||       |||   ||||||   ||
Db      525 uProGly--------------------GlnArgGlyGluArgGlyPheProGlyLeuPr 538

Qy      795 CCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGTCCGGGTCCCGG 736
             |                                ||||||      ||||||
Db      538 oGlyPro------------------SerGlyGluProGlyLysGlnGlyProSerGlyAl 552

Qy      735 AAGAGGTCAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCA-- 678
               |||:::|||   |||||||||                ||||||      |||
Db      552 aSerGlyGluArg---GlyProProGlyProMetGlyProProGlyLeuAlaGlyProPr 571

Qy      677 ----GCAGCCCCACGCCCAGGATGAGCAGTGCCAGCTCCAGGGGCC---TGGGATCCGGG 625
               :::   |||   |||                ||||||||||||       |||||
Db      571 oGlyGluSerGlyArgGluGly------------AlaProGlyAlaGluGlySerProGl 587

Qy      624 CACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCC 565
             |             ::::::                |||   |||   ||||||   |||
Db      587 yArgAspGlySerProGlyAlaLysGlyAspArgGlyGluThrGlyProAlaGlyAlaPr 607

Qy      564 AAGGACAGTGCCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTG 505
             |
Db      607 oGlyProPro------------------------------------------------- 610

Qy      504 GCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCATG 445
                         ||||||          |||||||||                ||||||
Db      611 ----------GlyAlaProGlyAlaProGlyProVal---------GlyProAlaGlyLy 624

Qy      444 GTCA------------TGAACTTCTCCTCTACCCCCACTT---CCAGCAGCAGAGGCGGC 400
             |||          |||   |||   |||:::   |||        |||
Db      624 sSerGlyAspArgGlyGluThrGlyProAlaGlyProIleGlyProValGlyProAlaGl 644

Qy      399 ACATAGGTGATGCCTGCGGCCAAACACACCTCCAGGCCAAAGG----------------- 357
               |||          |||      |||   |||   :::
Db      644 yAlaArgGlyProAlaGlyProGlnGlyProArgGly***LysGly***ThrGly***** 664

Qy      356 -------------TTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGT---GCCGCAGC 313
                         :::   |||                :::||||||    |||
Db      664 *Gly***ArgGlyIleLysGlyHisArgGlyPheSerGlyLeuGlnGlyProProGlyPr 684

Qy      312 AGGCGGCTCACCCACAGCCTCTGGACCATAGTGGG------------------------- 278
             |||||||||          |||||| ||||||
Db      684 oProGlySerProGlyGluGlnGlyPro-SerGlyAlaSerGlyProAlaGlyProArgG 704

Qy      277 -----CCAGGCGGGTAGGGCTCAGGGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCTCG 223
                  ||||||          ||||||   |||:::   |||                |||
```

```
Db        704 lyProProGlySerAlaGlySerProGlyLysAspGly-----------------LeuA 718

Qy        222 GCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAACTGCCTAGGAATCAGC 163
                        |||            |||                ||||||  :::
Db        718 snGlyLeuProGlyProIleGlyProProGly---------------ProArgGlyArgT 733

Qy        162 CAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGCTTCTCAGCCC 113
                 |||                    |||||||||||     |||
Db        733 hrGly--------------------AspAlaGlyProAlaGlyProPro 742

RESULT 30
B41182
collagen alpha 1(II) chain precursor (long splice form) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 16-Jul-1999
C;Accession: B41182
R;Metsaeranta, M.; Toman, D.; de Crombrugghe, B.; Vuorio, E.
J. Biol. Chem. 266, 16862-16869, 1991
A;Title: Mouse type II collagen gene. Complete nucleotide sequence, exon structure, and
A;Reference number: A41182; MUID:91358489; PMID:1885613
A;Accession: B41182
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-1487 <MET>
A;Cross-references: GB:M65161
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
C;Keywords: alternative splicing; coiled coil; extracellular matrix; glycoprotein; trime
F;33-91/Domain: von Willebrand factor type C repeat homology <VWC>
F;1259-1487/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Alignment Scores:
Pred. No.:              7.65e-12       Length:         1487
Score:                  321.00         Matches:        251
Percent Similarity:     30.67%         Conservative:   52
Best Local Similarity:  25.40%         Mismatches:     351
Query Match:            5.16%          Indels:         334
DB:                     2              Gaps:           51

US-09-759-143-110 (1-3410) x B41182 (1-1487)

Qy       2557 GGCCACATCCTGATAAAAGGTAAGAGGGGGGTGGATCAGCAAAAAGACAGTGCTGTGGGC 2498
               |||     |||    |||:::|||                  ||||||    |||
Db         32 GlySerCysLeuGlnAsnGlyGlnArg-----------TyrLysAspLysAspValTrp 47

Qy       2497 TGAGGGGACCTGGTTCTTGTGTGT---TGCCCCTCAGGACTCTTCCCCTACAAATAAGTC 2441
                                    :::|||  |||   :::|||                 :::
Db         48 LysProSerSerCysArgIleCysValCysAspThrGlyAsnValLeuCysAspAspIle 67

Qy       2440 ATATGTTCAAATCCCATGGAGGAGTGTTTCATCCTAGAAACTCCCATGCAAGAGCTACAT 2381
              ||||||   :::|||        :::|||            |||   |||
Db         68 IleCysGluAspPro------AspCysLeuAsnProGluIleProPhe----------- 81

Qy       2380 TAAACGAAGCTGCAGGTTAAGGGGCTTAGAGATGGGAAACCAGGTGACTGAGTTTATTCA 2321
                                                      |||:::
Db         82 ------------------------------------------GlyGluCysCysProIle 87

Qy       2320 GCTCCCAAAAACCCTTCTCTAGGTGTGTCTCAACTAGGAGGCTAGCTGTTAACCCTGAGC 2261
                |||   :::   :::          :::||||||
Db         88 CysProAlaAspLeuAlaThrAlaSerGlyLysLeuGly-------------------- 100

Qy       2260 CTGGGTAATCCACCTGCAGAGTCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCT 2201
                                                       |||    |||
Db        101 ------------------------------------------ProLysGly------ 103

Qy       2200 GTATAAGTCCAGACTGAAACCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGG 2141
                                                      |||   |||   |||   |||
Db        104 ------------------------------------GlnLysGlyGluProGlyAspIle 111

Qy       2140 GGAGAGAGGAGAGGGACGCCCCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAG 2081
                 :::       |||          ||||||      |||
Db        112 ArgAspIleIleGlyProArgGlyProProGlyProGln---------------GlyPro 126

Qy       2080 GGTGGCAGCAGAGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAG 2021
                |||    :::   |||                    ::::::      :::   |||
Db        127 AlaGlyGluGlnGlyProArgGlyAspArgGlyAspLysGlyGluLysGlyAlaProGly 146

Qy       2020 CCCCATGGGGCTAACAGGAGCGGGGAGCTGGGA-------------------------- 1988
               |||   |||         |||    ||||||   |||
Db        147 ProArgGly------ArgAspGlyGluProGlyThrProGlyAsnProGlyProAlaGly 164

Qy       1987 ---CCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTT 1931
                 |||            |||||||||      ||||||:::|||
Db        165 ProProGlyProProGlyProProGlyLeuSerAlaGlyAsnPhe-------------- 179

Qy       1930 GGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAG 1871
                                               :::::::::     |||          :::
Db        180 -----------------------------AlaAlaGlnMetAlaGlyGlyTyrAspGlu 189

Qy       1870 GCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCAT 1811
                   |||          :::|||  |||          |||        |||          |||
Db        190 LysAlaGlyGlyAlaGlnMetGly-Val-MetGlnGlyProMetGlyProMetGlyPro- 208

Qy       1810 AAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCACTATCCA---GGATGGCGAGGT- 1755
                        |||||||||||     ||||||           |||    |||::::::|||
Db        209 --------ArgGlyProProGlyProAlaGlyAlaProGlyProGlnGlyPheGlnGlyA 226

Qy       1754 ----------CCAGGCAGATGCCCCGGCCCGGAACCACCCTGG---CCTCGGTGGGCTC 1709
                         ||||||            ||||||          |||          ||||||   |||
Db        226 snProGlyGluProGlyGlu---ProGlyValSerGlyProMetGlyProArg--GlyPr 244

Qy       1708 ACCCACCACCACACGTACGGAGACATCACAGGCAGAGGCCCCGCAGAGCGCGGGTGGAGG 1649
               |||                              ||||||    :::             ||
Db        244 oProGlyProAlaGlyLysProGlyAspAspGlyGluAlaGlyLysProGlyLysSerGl 264

Qy       1648 T----------------------GGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCC 1613
              |                         |||:::|||      |||        |||        ||
Db        264 yGluArgGlyLeuProGlyProGlnGlyAlaArgGlyPheProGlyThrProGlyLeuPr 284

Qy       1612 ATTAGGGAAG---------GGAGCTCCAGGCTTAGGGCCTGGCAGGAAGCTGGTCATCAG 1562
              |    |||             |||    |||||||||          |||     |||          :::|
Db        284 oGlyValLysGlyHisArgGlyTyrProGlyLeu---AspGlyAlaLys-GlyGluAlaG 303

Qy       1561 GCTGTCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCTT 1502
              ||   |||                            |||:::          |||            |
Db        303 lyAlaPro--------------------------GlyVal---LysGlyGluSerGlyS 313

Qy       1501 CTCCCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAA 1442
              ||||||||      |||    |||                 |||          |||    ||| |||
Db        313 erProGlyGluAsnGlySerPro------------GlyProMetGlyProArg-GlyLeu 328

Qy       1441 CCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACC 1382
              |||||||||   |||   :::         |||             |||
Db        329 ProGlyGluArgGly--ArgThrGlyProAlaGlyAlaAlaGlyAlaArgGlyAsnAspG 348

Qy       1381 GGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCTG 1322
                ||||||
Db        348 lyGlnPro---------------------------------------------------G 351

Qy       1321 CACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACAGCC- 1263
                ||||||||||||                     ||||||
Db        351 lyProAlaGlyPro---------------ProGlyProValGlyProAlaGlyGlyProG 366

Qy       1262 -----CCAGGCTGCCCATCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGG 1208
                    ||||||   |||          :::               :::|||    |||      |||
Db        366 lyPheProGlyAlaProGlyAlaLysGlyGluAlaGlyProThrGlyAlaArgGlyProG 386

Qy       1207 CTCAGCTCTGGGCACGCCCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACAGCGT 1148
                                |||:::            |||    ||||||
Db        386 lu---------GlyAlaGlnGlySerArgGlyGluProGlyAsnProGlySerProGly- 402

Qy       1147 GAAGGTCATGAGTGCCATCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGTGCG 1088
                            |||:::         :::          |||              |||
Db        403 --------------ProAlaGlyAlaSerGlyAsnProGlyThrAspGlyIlePro---G 417

Qy       1087 GGGCATGCGGCAGCACAGCTGGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCCGGAAAGC 1028
              |||||    ||||||   |||||||||    |||     ||||||    ||||||          |
```

```
Db      417 lyAlaLysGlySer---AlaGlyAlaProGlyIleAlaGlyAlaProGlyPheProGlyP 436

Qy     1027 CAAGCGGGCCCGGCATGGACAGCAGTGGGGCGACA------------------------ 993
            ||   ||||||                        |||||||||
Db      436 roArgGlyPro---ProGlyProGlnGlyAlaThrGlyProLeuGlyProLysGlyGlnA 455

Qy      992 ---------------------------------AGGAGGGGGCCGACAGCCCTTCTGCTGG 965
                                             :::||||||
Db      455 laGlyGluProGlyIleAlaGlyPheLysGlyAspGlnGlyProLys-----------G 471

Qy      964 CTCGGTGGGGCCCAGCGCTGCCTCCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAG 905
                        |||||||||  |||       |||                     |
Db      471 lyGluThrGlyProAlaGlyProGlnGlyAlaPro----------------------G 483

Qy      904 GAAGATGAGGGTGAGCAGGCCAAAGAGGCACT--------------------------- 873
            ||       |||      |||:::||||||
Db      483 lyProAlaGlyGluGluGlyLysArgGlyAlaArgGlyGluProGlyGlyAlaGlyProI 503

Qy      873 ----------------------------------------------------------- 873

Db      503 leGlyProProGlyGluArgGlyAlaProGlyAsnArgGlyPheProGlyGlnAspGlyL 523

Qy      872 --------CCTCCTGGGTGCCCAGGT---AGGGGGCCAGGGCACTGGTGTCCCAGTCAAT 824
                    |||   |||   ||||||   |||||||||                :::
Db      523 euAlaGlyProLysGlyAlaProGlyGluArgGlyProSerGlyLeuAlaGlyProLysG 543

Qy      823 GGCAGGCAGGA------GGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCATAGAC 770
                  |||       |||   ||||||   |||                |||
Db      543 lyAlaAsnGlyAspProGlyArgProGlyGluProGlyLeuProGlyAlaArgGlyLeuT 563

Qy      769 AGAGTAGGCCTGGCGACAGTGGTCCGGGTCCCGGAAGAGGTCAGAGAGCAGGGCCTCCAG 710
                   |||||||||||:::||||||         |||:::         :::|||   |||
Db      563 hrGlyArgProGlyAspAlaGlyPro---GlnGlyLysValGlyProSerGlyAlaProG 582

Qy      709 TGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGC---------------- 666
              |||        ||||||           |||          |||
Db      582 lyGluAspGlyArgProGlyProProGlyProGlnGlyAlaArgGlyGlnProGlyValM 602

Qy      665 --------CCAGGATGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGCACAGCAGCCC 614
                    ||||||         |||          :::||||||    |||
Db      602 etGlyPheProGly---------ProLysGlyAlaAsnGlyGluProGlyLysAlaGly- 618

Qy      613 TGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCA------AGGA 560
                                          |||:::|||   ||||||   |||      |||
Db      619 -------------------------GluLysGlyLeuAlaGlyAlaProGlyLeuArgG 630

Qy      559 CAGTGCCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGA 500
              :::|||                                           |||
Db      630 lyLeuPro--------------------------------------GlyLysAspGlyG 637

Qy      499 GCCTAGGAGCGGGACACAGACCAGGCCCAGCA--------------------CTGGACC 461
                ||||||         ||||||||:::
Db      637 luThrGlyAlaAlaGlyProProGlyProSerGlyProAlaGlyGluArgGlyGluGlnG 657

Qy      460 AATGCCCAGCACCATGGTCATGAACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGG 401
                |||   |||             |||           ||||||
Db      657 lyAlaProGlyPro------------SerGlyPheGlnGlyLeuProGly---------- 669

Qy      400 CACATAGGTGATGCCTGCGGCCAAACACACCTCCAGGC------------CAAAGGTTAG 353
                                 |||   |||||||||          |||
Db      670 --------------------ProProGlyProProGlyGluGlyGlyLysGlnGlyAspG 683

Qy      352 CAGGTTGACCAGCAAGAGCTGGGCTTTCCGGTGCCGCAGCA------------------ 312
               |||   |||      ||||||      |||
Db      683 lnGlyIleProGlyGluAlaGlyAlaProGlyLeuValGlyProArgGlyGluArgGlyP 703

Qy      311 -----------GGCGGCTCACCCACAGCCTCTGGACCATAGTGGGCCAGGCGGG---TAG 266
                       |||||||||    |||    |||        ||||||   |||
Db      703 heProGlyGluArgGlySerProGlyAlaGlnGly-LeuGlnGlyProArgGlyLeuPro 722

Qy      265 GGCTCAGGGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCAGAAGCT 206
            |||:::      |||      |||   |||:::                     |||
Db      723 GlyThrProGlyThrAspGly---ProLysGlyAlaAlaGlyProAspGlyProProGly 741

Qy      205 GCGGCCTCTCCTCCTTGCTGCCGCCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGC 146
            |||      ||||||      :::   :::|||      :::     |||
Db      742 AlaGlnGlyProProGlyLeuGlnGlyMetProGlyGluArgGlyAlaAlaGlyIleAla 761

Qy      145 CAGCCCTTTGGT--------------------GCCGGTCCAGCTTCTCAGCCC------ 113
               |||   |||                      ||||||        |||
Db      762 GlyProLysGlyAspArgGlyAspValGlyGluLysGlyProGluGlyAlaProGlyLys 781

Qy      112 -----------------------ATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTGG 77
                                  :::       |||                :::||
Db      782 AspGlyGlyArgGlyLeuThrGlyProIleGly-ProProGlyProAlaGlyAlaAsnGl 801

Qy       76 GGACACGTCTCATCACTCAGATCCTGGC------------CGAGGCGCGCGGCTGTCACC 29
            |:::                  ||||||            |||||||||   :::||
Db      801 yGluLysGlyGluValGlyProProGlyProSerGlySerThrGlyAlaArgGlyAlaPr 821

Qy       28 CGGA 25
            ||||
Db      821 oGly 822
```

RESULT 31
S48789
sucrose transport protein - common tobacco
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 26-May-2000
C;Accession: S48789
R;Buerkle, X.Y.Z.; Frommer, W.B.
submitted to the EMBL Data Library, October 1994
A;Reference number: S48787
A;Accession: S48789
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-507 <BUE>
A;Cross-references: EMBL:X82276; NID:g575350; PID:g575351
C;Superfamily: common tobacco sucrose transport protein

Alignment Scores:

Pred. No.:	9.43e-12	Length:	507
Score:	319.50	Matches:	126
Percent Similarity:	41.04%	Conservative:	87
Best Local Similarity:	24.28%	Mismatches:	207
Query Match:	4.98%	Indels:	99
DB:	2	Gaps:	18

US-09-759-143-110 (1-3410) x S48789 (1-507)

```
Qy      278 CCCACTATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTC 337
             |||         :::||||||                               ::::::
Db       18 ProLeuProLysSerLysLeuTrp-----------------------------LysIle 27

Qy      338 TTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTAT 397
            ::::::|||        :::      |||::::::          |||      ::::::
Db       28 IleMetValAlaSerIleAlaAlaGlyValGlnPheGlyTrpAlaLeuGlnLeuSerLeu 47

Qy      398 GTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGC 457
            :::   |||   :::      :::|||:::         ||||||   :::   :::
Db       48 LeuThrProTyrValGlnLeuLeuGlyIleProHisLysPheAlaSerPheIleTrpLeu 67

Qy      458 ATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGG 517
               ||||||:::   |||::::::         |||:::::::::|||      ||||||:::
Db       68 CysGlyProIleSerGlyMetIleValGlnProValValGlyTyrTyrSerAspAsnCys 87

Qy      518 CGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTG 577
                  |||:::||||||||||||   ||||||   |||   :::      :::   :::
Db       88 SerSerArgPheGlyArgArgArgGlyPheIleAlaAlaGlyAlaAlaLeuValThrIle 107

Qy      578 AGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC--- 634
            ::::::|||||||||        |||         |||   |||       ||||||
Db      108 AlaValPheLeuIleGlyPheAlaAlaAspLeu---GlyHisAlaThrGlyAspProLeu 126

Qy      635 ------------AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGAC 682
                        :::|||      :::|||:::      ::::::|||      :::||||||
```

```
Db        127 GlyLysGlySerLysProArgAlaIleAlaValPheValValGlyPheTrpIleLeuAsp 146

Qy        683 TTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGAC 742
                          :::            |||       ||||||||:::||||||
Db        147 ValAlaAsnAsnMetLeuGlnGlyProCysArgAlaLeuLeuAlaAspLeuSerGlyGly 166

Qy        743 CCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGC 802
                          |||    :::    :::    ::::::|||    :::::::::|||
Db        167 LysAlaArgMetArgThrSerAsnAlaPhePheSerPhePheMetAlaValGlyAsnVal 186

Qy        803 CTGGGCTACCTCCTGCCTGCCATTGAC------------------------TGGGACACC 838
              |||||||||          :::                                  |||
Db        187 LeuGlyTyrAlaAlaGlySerTyrSerArgLeuCysLysIlePheProPheSerLysThr 206

Qy        839 AGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT---GGCCTGCTCACC 895
                 |||       |||              :::   |||   |||          :::
Db        207 ProAlaCysAspIleTyrCysAlaAsnLeuLysSerCysPhePheIleAlaValPheLeu 226

Qy        896 CTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGC 955
              |||:::   ||||||   :::|||   |||      |||       |||      |||
Db        227 LeuLeuSerLeuThrIleLeuAlaLeuThr------ValValArgGluAsnGluLeuPro 244

Qy        956 CCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGC 1015
                    :::      |||                  :::
Db        245 GluLysAspGluHisGlu------------------IleAspGluLysAlaGlyAlaArg 258

Qy       1016 CGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGC 1075
              :::::::::::    |||              |||   :::          |||      |||
Db        259 LysSerLysValProPhe------PheGlyGluIlePheGlyAlaLeuLysAspLeu--- 275

Qy       1076 TGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCA 1135
                       |||||    :::      |||     :::          :::|||:::|||
Db        276 ---------ProArgProMetTrpIleLeuLeuLeuValThrSerLeuAsnTrpIleAla 292

Qy       1136 CTCATGACCTTCACGCTGTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTG 1195
                       |||   |||:::   ||||||::::::   :::   :::|||   |||
Db        293 ArgPheProPhePheLeuTyrAspThrAspTrpMetAlaLysGluValTyrGlyGly--- 311

Qy       1196 CCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGC 1255
                 :::         |||                |||   ||||||   ||||||      |||
Db        312 ---LysValGlyAspGly------------ArgLeuTyrAspLeuGlyValHisAlaGly 326

Qy       1256 AGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGG 1315
              :::|||||||||   |||                 :::            ||||||   ::::::
Db        327 AlaLeuGlyLeuLeuLeuAsnSerValValLeuGlyPheMetSerLeuSerValGluPhe 346

Qy       1316 CTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTG 1375
              |||    ::::::    |||                  |||       :::    |||    :::
Db        347 LeuGlyLysLysIleGly---GlyValLysArgLeuTrpGlyIleLeuAsnPheValLeu 365

Qy       1376 GCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTG------------------ 1417
                       |||   |||::::::   :::|||   |||
Db        366 ---------AlaValCysMetAlaLeuThrValLeuValThrLysMetAlaGluLysSer 382

Qy       1418 ---------------------ACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCC 1456
                                   |||      |||      :::|||               |||
Db        383 ArgGlnTyrAspAlaHisGlyThrLeuMetAlaProThrSerGlyValLysIleGlyAla 402

Qy       1457 CTGCAGATC-------------------------------------CTGCCCTACACA 1477
              |||   :::                                     :::|||:::
Db        403 LeuThrLeuPheAlaValLeuGlyIleProLeuAlaValThrPheSerValProPheAla 422

Qy       1478 CTGGCCTCCCTCTACCACCGGGAGAAG--------------------------------- 1504
              |||||||||::::::
Db        423 LeuAlaSerIlePheSerSerAsnAlaGlySerGlyGlnGlyLeuSerLeuGlyValLeu 442

Qy       1505 ------CAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGAC 1558
                    :::   :::|||:::                  |||             :::|||
Db        443 AsnLeuAlaIleValValProGlnMetLeuValSerIleAlaGlyGlyProTrpAspAsp 462

Qy       1559 AGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACAC 1618
                                   |||||||||                  |||      |||
Db        463 LeuPheGlyGlyGlyAsnLeuProGly------------------PheIleValGlyAla 476

Qy       1619 GTGGGTGCTGGAGGCAGTGGC---------------CTGCTCCCACCTCCACCCGCG 1660
                 |||   |||      ||||||               :::||||||   |||||||||
Db        477 ValAlaAlaAlaAlaSerGlyIleLeuAlaLeuThrMetLeuProSerProProAla 495
```

RESULT 32
A41182
collagen alpha 1(II) chain precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 13-Aug-1999
C;Accession: A41182; A44885
R;Metsaeranta, M.; Toman, D.; de Crombrugghe, B.; Vuorio, E.
J. Biol. Chem. 266, 16862-16869, 1991
A;Title: Mouse type II collagen gene. Complete nucleotide sequence, exon structure, a
A;Reference number: A41182; MUID:91358489; PMID:1885613
A;Accession: A41182
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-1419 <MET>
A;Cross-references: GB:M65161
R;Cheah, K.S.; Lau, E.T.; Au, P.K.; Tam, P.P.
Development 111, 945-953, 1991
A;Title: Expression of the mouse alpha 1(II) collagen gene is not restricted to carti
A;Reference number: A44885; MUID:91347939; PMID:1879363
A;Accession: A44885
A;Molecule type: DNA
A;Residues: 1-28 <CHE>
A;Cross-references: GB:S63190; NID:g234368; PIDN:AAB19627.1; PID:g234369
A;Note: sequence extracted from NCBI backbone (NCBIN:63190, NCBIP:63192)
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo
C;Keywords: alternative splicing; coiled coil; extracellular matrix; glycoprotein; tr
F;1191-1419/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Alignment Scores:

Pred. No.:	1.08e-11	Length:	1419
Score:	318.50	Matches:	229
Percent Similarity:	31.66%	Conservative:	42
Best Local Similarity:	26.75%	Mismatches:	312
Query Match:	5.12%	Indels:	273
DB:	2	Gaps:	45

US-09-759-143-110 (1-3410) x A41182 (1-1419)

```
Qy       2164 CAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCCCCAGCCCCCAGCTGTG 2105
                 |||   |||   |||   |||         :::      |||          ||||||
Db         36 GlnLysGlyGluProGlyAspIleArgAspIleIleGlyProArgGlyProProGlyPro 55

Qy       2104 CAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGCAAC 2045
              |||                          |||   :::   |||                   :::
Db         56 Gln---------------GlyProAlaGlyGluGlnGlyProArgGlyAspArgGlyAsp 70

Qy       2044 AGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGGA--- 1988
              :::      :::   |||         |||   |||         |||   ||||||   |||
Db         71 LysGlyGluLysGlyAlaProGlyProArgGly------ArgAspGlyGluProGlyThr 88

Qy       1987 --------------------------CCCAGTGAGGCAGGCCCTCCACCCCAATGTGCT 1955
                                        |||      |||||||||             |||
Db         89 ProGlyAsnProGlyProAlaGlyProProGlyProProGlyProProGlyLeuSerAla 108

Qy       1954 GGAAGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGC 1895
              |||:::|||                                                  :::
Db        109 GlyAsnPhe------------------------------------------AlaAla 113

Qy       1894 AAAGTAAATGGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTG 1835
              ::::::   |||          :::          |||         :::|||   |||      |
Db        114 GlnMetAlaGlyGlyTyrAspGluLysAlaGlyGlyAlaGlnMetGly-Val-MetGlnG 133

Qy       1834 GCTGAGCTGGACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGC 1775
              ||      |||         |||               |||||||||||   ||||||
Db        133 lyProMetGlyProMetGlyPro---------ArgGlyProProGlyProAlaGlyAlaP 150

Qy       1774 ACTATCCA---GGATGGCGAGGT-----------CCAGGCAGATGCCCCGGCCCGGAAC 1730
                    |||   |||::::::|||                ||||||      ||||||
```

```
Db    150 roGlyProGlnGlyPheGlnGlyAsnProGlyGluProGlyGlu---ProGlyValSerG 169

Qy   1729 CACCCTGG---CCTCGGTGGGCTCACCCACCACCACGTACGGAGACATCACAGGCAGA 1673
              |||      ||||||  |||   |||                          ||
Db    169 lyProMetGlyProArg--GlyProProGlyProAlaGlyLysProGlyAspAspGlyGl 188

Qy   1672 GGCCCCGCAGAGCGCGGGTGGAGGT-----------------------GGGAGCAGGCC 1637
          ||||    :::               |||                   |||:::|||
Db    188 uAlaGlyLysProGlyLysSerGlyGluArgGlyLeuProGlyProGlnGlyAlaArgGl 208

Qy   1636 ACTGCCTCCAGCACCCACGTGTCCATTAGGGAAG---------GGAGCTCCAGGCTTAGG 1586
              |||     |||      |||      |||          |||   |||||||||
Db    208 yPheProGlyThrProGlyLeuProGlyValLysGlyHisArgGlyTyrProGlyLeu-- 227

Qy   1585 GCCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCTAGCACCTCCAGTGTCCCCTCG 1526
              |||   |||         :::|||   |||                          |
Db    228 -AspGlyAlaLys-GlyGluAlaGlyAlaPro--------------------------G 238

Qy   1525 GTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCAG 1466
          ||:::         |||          |||||||||      |||   |||           |
Db    238 lyVal---LysGlyGluSerGlySerProGlyGluAsnGlySerPro-----------G 253

Qy   1465 GATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCAC 1406
          ||        |||   ||| |||   |||||||||   |||   :::          |||
Db    253 lyProMetGlyProArg-GlyLeuProGlyGluArgGly--ArgThrGlyProAlaGlyA 272

Qy   1405 ACTGTGGGACAGGCATGTGGCACCGGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATA 1346
             |||                 ||||||
Db    272 laAlaGlyAlaArgGlyAsnAspGlyGlnPro--------------------------- 282

Qy   1345 GACTGCTCGAGTGCCGAATCGCTGCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGA 1286
                                  ||||||||||||               ||||||
Db    283 ----------------------GlyProAlaGlyPro---------------ProGlyP 290

Qy   1285 GATGGCGCACTGCAGGAACAGCC------CCAGGCTGCCCATCCGAACGCCTTCATCATA 1232
                         ||||||            ||||||    |||        :::
Db    290 roValGlyProAlaGlyGlyProGlyPheProGlyAlaProGlyAlaLysGlyGluAlaG 310

Qy   1231 GTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTACAGCCCCTCGCC 1172
              :::|||    |||   |||                |||:::                   |
Db    310 lyProThrGlyAlaArgGlyProGlu---------GlyAlaGlnGlySerArgGlyGluP 327

Qy   1171 CACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGCTGCACAGCTCAGC 1112
          ||   ||||||                      |||:::     :::                |
Db    327 roGlyAsnProGlySerProGly---------------ProAlaGlyAlaSerGlyAsnP 342

Qy   1111 CACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACAGCTGGTGCAGCCGGGGAAG 1052
          ||         |||         ||||||   ||||||   |||||||||   |||    |
Db    342 roGlyThrAspGlyIlePro---GlyAlaLysGlySer---AlaGlyAlaProGlyIleA 360

Qy   1051 CAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGGCCCGGCATGGACAGCAGTGGGGCGACA- 993
          |||||    ||||||         |||     ||||||               |||||||||
Db    360 laGlyAlaProGlyPheProGlyProArgGlyPro---ProGlyProGlnGlyAlaThrG 379

Qy    992 ---------------------------------------------------------AGGA 989
                                                                       :
Db    379 lyProLeuGlyProLysGlyGlnAlaGlyGluProGlyIleAlaGlyPheLysGlyAspG 399

Qy    988 GGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCTCCTCAGCCACCAG 929
          ::||||||                              |||||||||   |||         |||
Db    399 lnGlyProLys-----------GlyGluThrGlyProAlaGlyProGlnGlyAlaPro- 414

Qy    928 CAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGCACT---- 873
                                 |||      |||      |||:::||||||
Db    415 ----------------------GlyProAlaGlyGluGluGlyLysArgGlyAlaArgG 427

Qy    873 ------------------------------------------------------------ 873

Db    427 lyGluProGlyGlyAlaGlyProIleGlyProProGlyGluArgGlyAlaProGlyAsnA 447

Qy    872 ------------------------------CCTCCTGGGTGCCCAGGT---AGGGGGC 848
                                        |||   |||   ||||||    |||||||
Db    447 rgGlyPheProGlyGlnAspGlyLeuAlaGlyProLysGlyAlaProGlyGluArgGlyP 467

Qy    847 CAGGGCACTGGTGTCCCAGTCAATGGCAGGCAGGA------GGTAGCCCAGGCAGCCCCC 794
          ||                    :::            |||      |||   ||||||   |||
Db    467 roSerGlyLeuAlaGlyProLysGlyAlaAsnGlyAspProGlyArgProGlyGluProG 487

Qy    793 AAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGTCCGGGTCCCGGAA 734
                         |||                ||||||||||||:::||||||      |||:
Db    487 lyLeuProGlyAlaArgGlyLeuThrGlyArgProGlyAspAlaGlyPro---GlnGlyL 506

Qy    733 GAGGTCAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAG 674
          ::              :::|||   |||    |||             ||||||         |||
Db    506 ysValGlyProSerGlyAlaProGlyGluAspGlyArgProGlyProProGlyProGlnG 526

Qy    673 CCCCACGC-----------------------CCAGGATGAGCAGTGCCAGCTCCAGGGG 638
               |||                          ||||||       |||
Db    526 lyAlaArgGlyGlnProGlyValMetGlyPheProGly---------ProLysGlyAlaA 543

Qy    637 CCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCT 578
             :::||||||    |||                                 |||:::|||
Db    543 snGlyGluProGlyLysAlaGly--------------------------GluLysGlyL 554

Qy    577 CAGCAGGATGCCCA------AGGACAGTGCCCAGATGAAGGGCCGGCGGCGGCCATAGCG 524
             ||||||    |||        |||    :::|||
Db    554 euAlaGlyAlaProGlyLeuArgGlyLeuPro------------------------- 564

Qy    523 TCCACGCCAGTGGTCACTGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCA---- 468
                        |||                ||||||         |||||||||:::
Db    565 -----------GlyLysAspGlyGluThrGlyAlaAlaGlyProProGlyProSerGlyP 581

Qy    467 ----------------CTGGACCAATGCCCAGCACCATGGTCATGAACTTCTCCTCTAC 425
                                    |||   |||                  |||
Db    581 roAlaGlyGluArgGlyGluGlnGlyAlaProGlyPro-----------SerGlyPheG 597

Qy    424 CCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCCAAACACACCTCCAG 365
            ||||||                              |||      |||||||
Db    597 lnGlyLeuProGly----------------------------ProProGlyProProG 607

Qy    364 GC-----------CAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGTGCCG 317
          ||               |||      |||   |||      ||||||      |||
Db    607 lyGluGlyGlyLysGlnGlyAspGlnGlyIleProGlyGluAlaGlyAlaProGlyLeuV 627

Qy    316 CAGCA----------------------------GGCGGCTCACCCACAGCCTCTGGAC 287
                                          |||||||||   |||   |||
Db    627 alGlyProArgGlyGluArgGlyPheProGlyGluArgGlySerProGlyAlaGlnGly- 646

Qy    286 CATAGTGGGCCAGGCGGG---TAGGGCTCAGGGGGCCGTTCAGGCACTCCAGAACTGCTT 230
             ||||||   |||      |||:::      |||     |||   |||:::
Db    647 LeuGlnGlyProArgGlyLeuProGlyThrProGlyThrAspGly---ProLysGlyAla 665

Qy    229 CGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAACTGCCTAGG 170
                   |||      |||      ||||||         :::      :::|||
Db    666 AlaGlyProAspGlyProProGlyAlaGlnGlyProProGlyLeuGlnGlyMetProGly 685

Qy    169 AATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGT--------------------GCC 131
             :::      |||             |||   |||
Db    686 GluArgGlyAlaAlaGlyIleAlaGlyProLysGlyAspArgGlyAspValGlyGluLys 705

Qy    130 GGTCCAGCTTCTCAGCCC----------------------------ATGCTCAACACC 101
          ||||||      |||                                :::   ||
Db    706 GlyProGluGlyAlaProGlyLysAspGlyGlyArgGlyLeuThrGlyProIleGly-Pr 725

Qy    100 TGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTCATCACTCAGATCCTGGC-------- 49
          |                  :::|||:::                  ||||||
Db    725 oProGlyProAlaGlyAlaAsnGlyGluLysGlyGluValGlyProProGlyProSerGl 745

Qy     48 ----CGAGGCGCGCGGCTGTCACCCGGA 25
                ||||||||   :::||||||
Db    745 ySerThrGlyAlaArgGlyAlaProGly 754
```

RESULT 33
CGHU1E
collagen alpha 1(XI) chain precursor - human
N;Alternate names: procollagen alpha 1(XI) chain

```
C;Species: Homo sapiens (man)
C;Date: 31-Mar-1990 #sequence_revision 03-Oct-1995 #text_change 08-May-1998
C;Accession: A35239; A31795
R;Yoshioka, H.; Ramirez, F.
J. Biol. Chem. 265, 6423-6426, 1990
A;Title: Pro-alpha1(XI) collagen. Structure of the amino-terminal propeptide and express
A;Reference number: A35239; MUID:90202924; PMID:1690726
A;Accession: A35239
A;Molecule type: mRNA
A;Residues: 1-558 <YOS>
A;Cross-references: GB:J05407
R;Bernard, M.; Yoshioka, H.; Rodriguez, E.; van der Rest, M.; Kimura, T.; Ninomiya, Y.;
J. Biol. Chem. 263, 17159-17166, 1988
A;Title: Cloning and sequencing of pro-alpha1(XI) collagen cDNA demonstrates that type X
cartilagenous tissue.
A;Reference number: A92689; MUID:89034222; PMID:3182841
A;Accession: A31795
A;Molecule type: DNA; mRNA
A;Residues: 538-1806 <BER>
A;Cross-references: GB:J04177
A;Note: parts of this sequence were determined by protein sequencing
C;Comment: Prolines and lysines at the third position of the tripeptide repeating unit (
ed and subsequently O-glycosylated.
C;Genetics:
A;Gene: GDB:COL11A1; COLL6
A;Cross-references: GDB:120595; OMIM:120280
A;Map position: 1p21-1p21
A;Introns: 561/3; 579/3; 597/3; 615/3; 633/3; 648/3; 666/3; 681/3
A;Note: the list of introns is incomplete
C;Complex: type XI collagen may be a heterotrimer of two alpha 1(XI) chains and one alph
 3(XI) chain (see PIR:CGHU6C), initially linked by disulfide bonds among their carboxyl-
rmed with desmosine cross-links made from lysine and allysine residues
C;Function:
A;Description: structural component of extracellular fibrous polymer associated with cel
A;Note: may play a role in controlling the lateral growth of collagen II fibrils
C;Superfamily: collagen alpha 1(V) chain; fibrillar collagen carboxyl-terminal homology
C;Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxyproli
F;1-36/Domain: signal sequence #status predicted <SIG>
F;35-260/Domain: PARP-like #status predicted <PARP>
F;37-511/Domain: amino-terminal propeptide #status predicted <PRO>
F;512-1565/Product: collagen alpha 1(XI) chain #status predicted <MAT>
F;512-527/Region: amino-terminal nonhelical telopeptide
F;528-1542/Region: helical
F;1543-1565/Region: carboxyl-terminal nonhelical telopeptide
F;1566-1806/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F;1583-1805/Domain: fibrillar collagen carboxyl-terminal homology <FCC>
F;61-243,182-236/Disulfide bonds: #status predicted
F;505/Modified site: allysine (Lys) #status predicted
F;612,1452/Modified site: 5-hydroxylysine (Lys) #status predicted
F;612,1452/Binding site: carbohydrate (Lys) (covalent) #status predicted

Alignment Scores:
Pred. No.:              1.54e-11       Length:         1806
Score:                  316.00         Matches:        315
Percent Similarity:     31.07%         Conservative:   61
Best Local Similarity:  26.03%         Mismatches:     413
Query Match:            5.08%          Indels:         423
DB:                     1              Gaps:           71

US-09-759-143-110 (1-3410) x CGHU1E (1-1806)

Qy     3101 ACAAAGGCTTGGGAAACCGCACTTTGTGCTTCTGGT----------------------- 3066
               |||        ||||||:::::         |||
Db      409 ThrAspIleThrGluThrSerIleAsnGlyHisGlyAlaTyrGlyGluLysGlyGlnLys 428

Qy     3065 ------CCTGCAGTAGCT---CCAAACAGGGTTGTGGAGCTGGTGGGGAAAGTTGGGGGT 3015
                  |||||||||       |||      :::||||||               |||
Db      429 GlyGluProAlaValValGluProGlyMetLeuValGlu---------------GlyPro 443

Qy     3014 AGGGGAAAGTTGGGGGTAGGGGAAATTTTGGGCAGTGCCTTCATCAGCCCAGTCCTAGAG 2955
               |||        |||          |||:::|||             |||   |||:::
Db      444 ProGlyProAlaGlyProAlaGlyIleMetGly-----------ProProGlyLeuGln 459

Qy     2954 AGAGTAGAGGGGAGTGGAAGTGGGGGGAACCAGGCTGGGCCAAGAGAAGAGGGGTGGTTA 2895
                              |||      |||          |||    :::   |||
Db      460 --------------GlyProThrGlyProProGlyAspProGlyAspArgGlyProPro 474

Qy     2894 GGGAAGCCGTTGAGACCTGAAGCC--------------CCACCCTCTACCTTCCTTCAA 2850
            |||:::|||      |||    |||                ||||||   |||   |||
Db      475 GlyArgProGlyLeuProGlyAlaAspGlyLeuProGlyProProGlyThrMetLeuMet 494

Qy     2849 CACCCTAACCTTGGGTAACAGCATTTGGAATTATCATTTGGGATGAGTAGAATTTCCAAG 2790
               |||                                :::|||
Db      495 LeuPro----------------------PheArgTyrGlyGlyAsp----------- 502

Qy     2789 GTCCTGGGTTAGGCATTTTGGGGGGCCAGA--------------CCCCAGGAGAAGAAG 2745
                                     |||::::::                ||||||   :::
Db      503 --------------------GlySerLysGlyProThrIleSerAlaGlnGluAlaGln 515

Qy     2744 ATTCTGGCAATGATCAGCCCAATGACCAGCTATCTCAGGGGACCTGAT----------- 2697
                     |||::::::                ||||||||||||
Db      516 AlaGlnAlaIleLeuGlnGlnAlaArgIleAlaLeuArgGlyProProGlyProMetGly 535

Qy     2696 ---TGTTGGGGATCCCCCACCCTACCCAAATATTAGACACCAACACAGAAAAGCTAGCAA 2640
                   |||   |||   |||                           |||      ::
Db      536 LeuThr-GlyArgProGlyProValGlyGlyProGlySerSerGlyAlaLysGlyGluSe 555

Qy     2639 TGGATTCCCTTCTACTTTGTTAAATAAATAAGTTAAATATTTAAATGCCTGTGTCTCTGT 2580
            :|||   |||                                               ||
Db      555 rGlyAspProGlyPro------------------------GlnGlyProArgGlyVa 566

Qy     2579 GATGGCAACAGAAGGACCAACAGGCCACATCCTGATAAAAGG----------------- 2538
            |               |||||||||||      |||    ||||||
Db      566 lGlnGlyProProGlyProThrGly--LysProGlyLysArgGlyArgProGlyAlaAsp 585

Qy     2537 -----------------TAAGAGGGGGGTGGATCAGCAAAAAGACAGTGCTGTGGGCTG 2496
                                 |||      |||                |||         ||||||
Db      586 GlyGlyArgGlyMetProGlyGluProGlyAlaLysGlyAspArgGlyPheAspGlyLeu 605

Qy     2495 AGGGGA---CCTGGTTCTTGTGTGTTGCCCCTCAGGACTCTTCCCCTACAAATAAGTCAT 2439
               |||   ||||||
Db      606 ProGlyLeuProGlyAsp----------------------------------------- 611

Qy     2438 ATGTTCAAATCCCATGGAGGAGTGTTTCATCCTAGAAACTCCCATGCA----------- 2391
                  |||   |||   |||         |||:::
Db      612 ------LysGlyHisArgGlyGluArgGlyProGlnGlyProProGlyProProGlyAsp 629

Qy     2390 ----------------------------------AGAGCTACATTAAACGAAGCTGCA 2367
                                                  |||         ||||||
Db      630 AspGlyMetArgGlyGluAspGlyGluIleGlyProArgGlyLeuProGlyGluAlaGly 649

Qy     2366 GGTTAAGGGGCTTAGAGAT------GGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAA 2313
                :::||||||            |||   ||||||                  ||||||
Db      650 -ProArgGlyLeuLeuGlyProArgGlyThrProGly--------------AlaProGl 664

Qy     2312 AAACCCTTCTCTA---GGTGTGTCTCAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGG 2256
                |||   :::    ||||||                                   ||
Db      664 yGlnProGlyMetAlaGlyValAspGlyProProGly--------------ProLysGl 679

Qy     2255 TAAT-----------------------CCACCTGCAGAGTCCCCGCATTCCAGTGCATG 2220
            ||||                        ||||||    :::         :::
Db      679 yAsnMetGlyProGlnGlyGluProGlyProProGlyGlnGlnGlyAsnPro-------- 696

Qy     2219 GAGCCCTTCTGGCCTCCCT---GTATAAGTCCAGACTGAAACCCCCTTGGAAGGCCTCCA 2163
                |||   |||||||||         :::       :::   ||||||   :::
Db      697 -GlyProGlnGlyLeuProGlyPro-GlnGlyProIleGlyProProGlyGluLysGlyP 716

Qy     2162 GTCAG---------------GCAGCCCTA---------------------------- 2149
              |||               |||   |||
Db      716 roGlnGlyLysProGlyLeuAlaGlyLeuProGlyAlaAspGlyProProGlyHisProG 736

Qy     2148 -----------GAGACTGGGGAGAGAGGAGAG---GGACGCCCCAGCCCCCAGCTGTGCA 2103
                          ::::::|||||||:::|||        |||    |||    ||||||
Db      736 lyLysGluGlyGlnSerGlyGluLysGlyAlaLeuGlyProProGlyProGlnGlyProI 756

Qy     2102 GCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGCAACAG 2043
```

```
Db    756 leGly***ProGlyProArgGlyValLysGlyAlaAspGlyValArgGlyLeuLysGlyS 776

Qy   2042 AAACTGGC------GGCCAGCCCGGCAGCCCC---------------ATGGGGCTAACAG 2004

Db    776 erLysGlyGluLysGlyGluAspGlyPheProGlyPheLysGlyAspMetGlyLeuLysG 796

Qy   2003 GAGC-------------GGGGAGCTGGGACCC---------------------------- 1985

Db    796 lyAspArgGlyGluValGlyGlnIleGlyProArgGly***AspGlyProGluGlyProL 816

Qy   1984 --AGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGCC 1927

Db    816 ysGlyArgAlaGlyProThrGlyAspProGlyProSer------------------GlyG 830

Qy   1926 AAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAGAC-------- 1875

Db    830 lnAla-----------------------------------GlyGluLysGlyLysLeuG 838

Qy   1874 -----CCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTGGACAAT 1820

Db    838 lyValProGlyLeuProGlyTyrProGlyArgGln------------------------ 849

Qy   1819 GGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCACTATCCAGGATGG- 1761

Db    850 --------------------GlyProLysGlySerThrGly------PheProGlyPheP 861

Qy   1760 -------------------------CGAGGTCCAGGCAGATGCCCCGGCCCGGAA---- 1731

Db    861 roGlyAlaAsnGlyGluLysGlyAlaArgGlyValAlaGlyLysProGlyProArgGlyG 881

Qy   1730 --------CCACCCTGGCCTCGGTGGGCTCACCCACCACCACACGTACGGAGACATCACA 1679

Db    881 lnArgGlyProThrGlyProArg--GlySer---ArgGlyAlaArgGlyProThrGlyLy 899

Qy   1678 GGCAGAGGCCCCGCAGAGCGCGGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCAC 1619

Db    899 s---ProGlyProLysGlyThrSerGlyGlyAspGlyProProGlyProProGlyGluAr 918

Qy   1618 GTGTCCATTAGGGAAGGGAGCTCCAGGCTTAGGGCCTGGCAGGAAGCTGGTCATCAGGCT 1559

Db    918 gGlyProGlnGlyProGlnGlyProValGlyPheProGlyProLys-------------- 933

Qy   1558 GTCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTC 1499

Db    934 -------------GlyProProGlyProPro-Gly---ArgMetGlyCysProGlyHisP 948

Qy   1498 CCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCC 1439

Db    948 roGlyGlnArgGly-------------GluThrGly-PheGlnGlyLysThrGlyProPro 963

Qy   1438 GGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACCGGC 1379

Db    964 GlyProGlyGly--------------------ValValGlyProGlnGlyProThrGly 976

Qy   1378 AGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCTGCAC 1319

Db    977 Glu-ThrGly-----------------------------------Pro-IleGlyGluA 984

Qy   1318 CAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACAGCC---- 1263

Db    984 rgGlyTyrProGlyProProGlyProProGlyGluGlnGlyLeuProGlyAlaAlaGlyL 1004

Qy   1262 --CCAGGCTGCCCATCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTC 1205

Db   1004 ysGluGlyAlaLysGlyAspProGlyProGlnGlyIleSerGlyLysAspGlyProAlaG 1024

Qy   1204 AGCTCTGGGCACGCCCTGGTACA------------------------GCCCCT------- 1176

Db   1024 lyLeuArgGlyPheProGlyGluArgGlyLeuProGlyAlaGlnGlyAlaProGlyLeuL 1044

Qy   1175 -----------CGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCA 1127

Db   1044 ysGlyGlyGluGlyProGlnGlyPro--------------------------ProGlyP 1055

Qy   1126 GCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACAGCTG 1067

Db   1055 roValGlySerProGlyGluArgGlySerAlaGly------------------ThrAlaG 1069

Qy   1066 GTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGGCCCGGCATGGACA 1007

Db   1069 lyProIleGlyLeuArgGlyArgProGlyProGlnGlyProProGlyProAlaGlyGluL 1089

Qy   1006 GCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCGC 947

Db   1089 ysGlyAlaProGlyGluLysGlyProGln--------------------GlyProAla- 1101

Qy    946 TGCCTCCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAG 887

Db   1102 ----------------------------------------GlyArgAspGlyValGlnG 1108

Qy    886 GCCAAAGAGGCACTCCTCCTGGGTGCCCAGGT---AGGGGGCCAGGGCACTGGTGTCCCA 830

Db   1108 ly------------ProValGlyLeuProGlyProAlaGlyPro---------------- 1118

Qy    829 GTCAATGGCAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCATAGAC 770

Db   1119 --------------AlaGlySerProGlyGluAspGlyAsp---------LysGlyGluI 1131

Qy    769 AGAGTAGGCCTGGCGACAGTGGTCCGGGTCCCGGAAGAGGTCAGAGAGCAGGGCCTCCAG 710

Db   1131 leGlyGluProGlyGlnLysGlySerLysGlyGlyLysGlyGluAsn---GlyProPro- 1149

Qy    709 TGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGCCCAGGATGAGCAGTGC 650

Db   1150 --------GlyProProGlyLeuGlnGlyProValGly---AlaProGly---------I 1163

Qy    649 CAGCTCCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGAT 590

Db   1163 leAlaGlyGlyAspGlyGluProGly---------------------------------- 1171

Qy    589 GAGAAAGAGGCTCAGCAGGATGCCCAAGGACAGTGCCCAGATGAAGGGCCGGCGGCGGCC 530

Db   1172 --ProArgGly-GlnGlnGlyMetPheGlyGlnLysGlyAspGluGly---AlaArgGly 1189

Qy    529 ATAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAG 470

Db   1190 PheProGlyProPro--Gly---------ProIleGlyLeuGlnGlyLeuProGlyProP 1206

Qy    469 CACTGGACCAATGCCCA------------GCACCATGGTCATGAACTTCTCCTCTACCCC 422

Db   1206 roGlyGluLysGlyGluAsnGlyAspValGlyProTrpGly------ProProGlyProP 1224

Qy    421 CACTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCCAAACACACCTCCAGGCC 362

Db   1224 roGlyProArgGlyProGlnGlyProAsnGlyAlaAspGlyProGlnGlyProProGly- 1243

Qy    361 AAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGTGCCGCAGCAGGCGGCTCAC 302

Db   1244 -------------------SerValGlySerValGlyGlyValGlyGluLysGlyGluP 1257

Qy    301 CCACAGCCTCTGGACCATAGTGGG---CCAGGCGGGTAGGGCTCAGGGGGCCGTTCAGGC 245

Db   1257 roGlyGluAlaGly-AsnProGlyProProGlyGluAlaGlyValGlyGlyProLysGly 1276

Qy    244 ACTCCAGAACTGCTTCGTCTCGGCTCTGCT---CCAGAAGCTGCGGCCTCTCCTCCTTGC 188

Db   1277 ------GluArgGlyGluLysGlyGluAlaGlyProProGlyAlaAlaGlyProProGly 1294

Qy    187 TGCCGC---------------CAACTGCCTAGGAATCAGCCAGGCGCCCATTTCT----- 148

Db   1295 AlaLysGlyProProGlyAspAspGlyProLysGlyAsnProGly-ProValGlyPhePr 1314

Qy    147 -GCCAGCCCTTTGGTGCCGGTCCAGCTTCTCAGCCCATGCTCAACACCTGCTGCTGTGGG 89

Db   1314 oGlyAspProGlyProProGlyGluLeu---------------GlyProAlaGlyGlnAs 1329
```

```
Qy          88 GCACCTCAGTGGGGACACGTCTCATCACTCAGATCCTGGCCGA---------------- 46
                   :::   ||||||                 |||||||||:::
Db        1329 pGlyValGlyGlyAspLysGlyGluAspGlyAspProGlyGlnProGlyProProGlyPr 1349

Qy          45 ------------------------------GGCGCGCGGCTGTCACCCGGAGCCAGC 19
                                             |||   |||      |||||||||:::
Db        1349 oSerGlyGluAlaGlyProProGlyProProGlyLysArgGlyProProGlyAlaAla 1368
```

RESULT 34
S16366
collagen alpha 2(IV) chain precursor - pig roundworm
C;Species: Ascaris suum (pig roundworm)
C;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 13-Aug-1999
C;Accession: S16366
R;Pettitt, J.; Kingston, I.B.
J. Biol. Chem. 266, 16149-16156, 1991
A;Title: The complete primary structure of a nematode alpha-2(IV) collagen and the parti
A;Reference number: S16366; MUID:91340768; PMID:1714907
A;Accession: S16366
A;Molecule type: mRNA
A;Residues: 1-1763 <JBI>
A;Cross-references: GB:M67507; NID:g159648; PIDN:AAA18014.1; PID:g159649
C;Genetics:
A;Introns: 229/3; 266/3; 305/3; 360/3; 424/1; 489/1; 548/1; 656/3; 790/1; 891/1; 963/1;
C;Superfamily: collagen alpha 1(IV) chain
C;Keywords: alternative splicing; basement membrane; cell binding; coiled coil; disulfid
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-1763/Product: collagen alpha 2(IV) chain #status predicted <MAT>
F;27-42/Domain: non-collagenous NH1 #status predicted <NH1>
F;43-1529/Domain: collagenous #status predicted <COL>
F;197-199/Region: cell attachment (R-G-D) motif
F;1530-1763/Domain: carboxyl-terminal nonhelical, NC1 #status predicted <NC1>
F;1530-1638/Domain: repeat NC1 #status predicted <NC11>
F;1639-1763/Domain: repeat NC1 #status predicted <NC12>
F;31,34,39,41,536,539/Disulfide bonds: interchain #status predicted
F;126/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;1593-1599,1702-1709/Disulfide bonds: #status predicted

Alignment Scores:			
Pred. No.:	2.5e-11	Length:	1763
Score:	312.50	Matches:	315
Percent Similarity:	31.65%	Conservative:	60
Best Local Similarity:	26.58%	Mismatches:	430
Query Match:	4.87%	Indels:	381
DB:	2	Gaps:	68

US-09-759-143-110 (1-3410) x S16366 (1-1763)

```
Qy           2 GGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAG 61
                |||      |||      |||     ||||||           |||   |||   |||
Db         404 GlyProGlyLeuProGlyLeu---ProGlyLeuGluGlyLeuProGlyProLysGlyGlu 422

Qy          62 TGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAG------CATGGG 115
                        |||             |||   |||   ||||||:::          |||
Db         423 LysGlyAspSerGly---------IleProGlyAlaProGlyValGlnGlyProProGly 439

Qy         116 CTGAGAAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCT-----------GG 163
                |||          |||   ||||||       |||   ||| |||           ||
Db         440 LeuAlaGlyProProGlyAlaLysGly------GluProGly-ProArgGlyValAspGl 457

Qy         164 CTGATTCCTAGGCAGTTGGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCC 223
                |        :::             :::   |||               |||
Db         457 yGlnSerIleProGlyLeuProGlyLysAspGlyArgProGlyLeuAspGlyLeuProGl 477

Qy         224 GAGACGAAGCAGTTCTGGAGTGCCT------GAACGGCCCCCTGAG------------- 263
                |||:::              |||:::|||          ||||||
Db         477 yArgLysGlyGluMetGlyLeuProGlyValArgGlyProProGlyAspSerLeuAsnGl 497

Qy         264 ----CCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGG 319
                    |||   ||||||||||||     ||||||:::
Db         497 yLeuProGlyProProGlyPro-ArgGlyProGln-----------------------G 509

Qy         320 CACCGGAA--------------------AGCCCAGCTCTTGCTGGTCAACCTGCTAACC 358
                |||:::                      :::|||   |||   |||   |||
Db         509 lyProLysGlyTyrAspGlyArgAspGlyAlaProGlyLeuProGlyIlePro------- 526

Qy         359 TTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAA 418
                          |||   |||   |||         ||||||   |||
Db         527 ----------GlyProLysGlyAspArgGlyGlyThrCysAlaPheCysAla------H 541

Qy         419 GTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCT- 477
                |||   :::||||||      |||        ||||||:::           ||||||
Db         541 isGlyAlaLysGlyGluLysGlyAsp------AlaGlyTyrAlaGlyLeuProGlyProG 559

Qy         478 -------------------------------------------------------GGTC 481
                                                                       ||||
Db         559 lnGlyGluArgGlyLeuProGlyIleProGlyAlaThrGlyAlaProGlyAspAspGlyL 579

Qy         482 TGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGG 541
                ||     ||||||                   |||                 ||||||
Db         579 euProGlyAlaProGlyArgProGlyProPro--------------GlyProProGlyG 594

Qy         542 CCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCC 601
                       |||       :::   |||      ||||||                      :
Db         594 lnAspGlyLeuProGlyLeuProGlyGlnLysGlyGluProThr---------------G 609

Qy         602 GGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATC 661
                :::|||       |||         ||||||                       |||    |
Db         609 lnLeuThrLeuArgProGlyProProGlyTyrProGlyGlnLysGlyGluThrGlyPheP 629

Qy         662 CTGGG---CGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCC 718
                |||||   ||||||         ||||||       ||||||            |||   |
Db         629 roGlyProArgGlyGlnGluGlyLeu------ProGlyLysProGlyIleValGlyAlaP 647

Qy         719 CTGCTCTCTGACCTCTTCCGGGA---CCCGGACCACTG---TCGCCAGGCCTACTCTGTC 772
                ||               |||||||||   |||||||||          |||||||||   |
Db         647 roGly---------LeuProGlyProProGlyProLysGlyGluProGlyLeuThrGlyL 664

Qy         773 TATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGG 832
                ||                          |||:::||||||||||||             |
Db         664 euProGluLysProGlyLysAspGlyIleProGlyLeuProGlyLeuLysGlyGluProG 684

Qy         833 GACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGA------GGAGTGCCTCTTTGGC 886
                |||:::     ||||||   :::||||||   ||||||      |||            |
Db         684 lyTyrGlyGlnProGly---MetProGlyLeuProGlyMetLysGlyAspAlaGlyLeuP 703

Qy         887 CTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCA 946
                ||     |||   ||||||                        |||      |||
Db         703 roGlyLeuProGlyLeuPro-----------------------GlyAlaValGlyProM 715

Qy         947 GCGCTGGGCCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGC 1006
                        |||         |||:::         ||||||               |||
Db         715 etGlyProProValProGluSerGlnLeuArgProGlyProProGlyLysAspGlyLeuP 735

Qy        1007 TGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCT---- 1062
                  :::|||||||||   |||         |||   |||   |||   :::|||
Db         735 roGlyLeuProGlyProLysGly------GluAlaGlyPheProGlyAlaProGlyLeuG 753

Qy        1063 --GCACCAGCTGTGCTGCCGCATGCCCCG-----------CACCCT-----------G 1096
                  ||||||   ||||||   |||                  :::|||           |
Db         753 lnGlyProAlaGlyLeuProGlyLeuProGlyMetLysGlyAsnProGlyLeuProGlyA 773

Qy        1097 CGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTT- 1155
                |||||   |||   |||                    |||   |||              :::
Db         773 laProGlyLeuAlaGly-----------LeuProGlyIleProGlyGluLysGlyIleA 789

Qy        1156 ----------------------TTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGC 1192
                                      |||   |||   :::|||           ||||||:
Db         789 laGlyLysProGlyLeuProGlyLeuThrGlyAlaLysGlyGluAlaGlyTyrProGlyG 809

Qy        1193 GT------GCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTT 1246
                ::                |||      ||||||   |||
Db         809 lnProGlyLeuProGlyProLysGlyGluProGlyProSerThrThrGlyProProGlyP 829

Qy        1247 CGGATGGGCAGCCTGGG---------------GCTGTTCCTGCAGTGCGCCATCTCCCTG 1291
```

```
                  |||  ||||||                   :::|||             |||||||
Db      829 roProGlyPheProGlyLeuLysGlyLysAspGlyIleProGlyAlaProGlyLeuProG 849

Qy     1292 GTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTG 1351
            |||||          |||      |||        :::           :::  |||
Db      849 lyLeuGluGlyGlnArgGlyLeuProGlyValProGlyGlnLysGlyGlu-IleGlyLeu 868

Qy     1352 GCCAGTGTGGCA---------GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCAC 1402
               :::|||                 ||||||  |||   |||         |||  |||
Db      869 ProGlyLeuAlaGlyAlaProGlyPheProGlyAlaLysGlyGluProGlyLeu-ProGl 888

Qy     1403 AGTG--------------------TGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGG 1441
              :::                        |||         ||||||   :::|||||
Db      888 yLeuProGlyLysGluGlyProGlnGlyPro------ProGlyGlnProGlyAlaProGl 906

Qy     1442 TTCACCTTCTCAGCCCTGCAGATCCTGCCCTA--------------------------- 1473
             |   |||   |||        |||        |||
Db      906 yPheProGlyGln-LysGlyAspGluGlyLeuProGlyLeuProGlyValSerGlyMetL 926

Qy     1474 -----CACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAA------A 1522
                      ||||||||||||      ||||||  ||||||   |||    |||
Db      926 ysGlyAspThrGlyLeuProGlyValProGlyLeuAlaGlyProProGlyGlnProGlyP 946

Qy     1523 TACCGAGGGGACACTGGA-----------------------GGTGCTAGCAGTGAGGAC 1558
              ||| |||        |||                       ||||||    |||
Db      946 hePro-GlyGlnLysGlyGlnProGlyPheProGlyValAlaGlyAlaLysGlyGlu--- 964

Qy     1559 AGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACAC 1618
                            :::   |||||||||       |||||||||||        |||
Db      965 ------------AlaGlyLeuProGly---LeuProGlyAlaProGlyGlnLysGlyGlu 979

Qy     1619 GTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCC 1678
               |||   |||               ||||||    ||||||  |||   :::   ||||||
Db      980 GlnGlyLeuAlaGly---------LeuProGlyIleProGlyMetLysGlyAlaProGly 996

Qy     1679 TGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGC--------------- 1722
                                  :::   |||   |||   :::
Db      997 --------------------IleProGlyAlaProGlyGlnAspGlyLeuProGlyLeu 1009

Qy     1723 ---------CAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGC---CATCCTGGA 1770
                       :::|||      |||         |||||||||   :::     ||||||
Db     1010 ProGlyValLysGlyAspArgGlyPheAsnGlyLeuProGlyGluLysGlyGluProGly 1029

Qy     1771 TAGTGCCTT--------------------CCTGCTGTCCCAGGTGGCCCCATCCCTGTT 1809
                                         |||    :::||||||     |||
Db     1030 ProAlaAlaArgAspGlyGluLysGlyGluProGlyLeuProGlyGlnPro--------- 1046

Qy     1810 TATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGG 1869
               ||||||  |||       |||                     |||:::
Db     1047 ---GlyLeuArgGlyProGlnGlyProPro---------------GlyLeuProGlyLeu 1060

Qy     1870 CCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGC 1929
            ||||||
Db     1061 ProGly------------------------------------------------------ 1062

Qy     1930 CAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTC 1989
               :::         |||               |||     |||||| :::
Db     1063 ---LeuLysGlyAspGluGlyGlnProGlyTyrGlyAlaProGlyLeu-MetGlyGluLy 1081

Qy     1990 CCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGC 2049
                ||||||                       ||||||||||||   |||
Db     1081 sGlyLeuPro-----------------GlyLeuProGlyLysPro-------------- 1090

Qy     2050 TGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAG 2109

Db     1090 ------------------------------------------------------------ 1090

Qy     2110 CTGGGGGCTGGGGCGTCCC---TCTCCTCTCTCCCCAGTCTCTA---GGGCTGCCTGACT 2163
                          |||||||||   :::|||   |||:::   |||   |||   |||   :
Db     1091 ---------GlyArgProGlyAlaPro-GlyProLysGlyLeuAspGlyAlaProGlyP 1107

Qy     2164 GGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATG 2223
            ::              ||||||   :::                   |||   |||   |||
Db     1107 heProGlyLeuLysGlyGluAla---------------GlyLeuPro---GlyAlaProG 1121

Qy     2224 CACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTC 2283
               |||        |||         |||||||||     :::        |||      |
Db     1121 lyLeuProGlyGlnAspGlyLeuProGlyLeuProGlyGlnLysGlyGluSerGlyPheP 1141

Qy     2284 CTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTT 2343
            ||                   :::    |||    :::|||                 ||||||
Db     1141 roGly------------GlnProGlyLeuValGly---------------ProProGlyL 1152

Qy     2344 TCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTT 2403
             |||                                                  ||||||
Db     1152 euProGlyLys------------------------------------------MetGlyAlaP 1159

Qy     2404 CTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAG---- 2459
                   ::::::               |||         :::        |||    |||:::
Db     1159 roGlyIleArg---GlyGluLysGlyAspAlaGlyLeuProGlyLeuProGlyGluArgG 1178

Qy     2460 --------AGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGT 2511
                                ||||||                       |||
Db     1178 lyLeuAspGlyLeuProGlyGlnLysGlyGluAlaGlyPhePro--------------- 1192

Qy     2512 CTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTG 2571
                               |||   ||||||                  |||||||||||
Db     1193 ----------GlyAlaProGlyLeuPro------------GlyProValGlyProLysG 1205

Qy     2572 TTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAG 2631
              :::                                                  |||     ::::::|
Db     1205 lySerAlaGlyAlaProGly-----------------------PheProGlyLeuLysG 1217

Qy     2632 GGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGTAGGGTGGGGGATCCC 2691
              ||   |||   |||        |||                              |||      |
Db     1217 lyGluProGlyLeuProGlyLeu--------------------------GluGlyGlnP 1228

Qy     2692 CAACAATCAGGTCC--------------------------------------------- 2705
              ||      |||
Db     1228 roGlyProArgGlyMetLysGlyGluAlaGlyLeuProGlyAlaProGlyArgAspGlyL 1248

Qy     2706 --CCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTG 2763
               |||     :::     |||          :::            |||      |||
Db     1248 euProGlyLeuProGlyMetLysGlyGlu------AlaGlyLeuProGlyLeuProGlyG 1266

Qy     2764 GCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCC 2823
               |||   |||      :::|||          ||||||:::         :::|||
Db     1266 lnProGlyLysSerIleThrGlyProLysGlyAsnAlaGly-LeuProGlyLeuProGly 1285

Qy     2824 AAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGT-------------- 2868
             |||             |||   |||             :::      |||
Db     1286 LysAspGlyLeuProGlyLeuProGlyLeuLysGlyGluProGlyLysProGlyTyrAla 1305

Qy     2869 -----------------------------------------------------GGGGCT 2874
                                                                 |||
Db     1306 GlyAlaAlaGlyIleLysGlyGluProGlyLeuProGlyIleProGlyAlaLysGlyGlu 1325

Qy     2875 TCAGGTCTCAACGGCTTCCCTAACCAC----------------CCCTCTTCTCTTGGC 2916
               ||||||:::|||    |||                          |||        |||
Db     1326 ProGlyLeuSerGlyIleProGlyLysArgGlyAsnAspGlyIleProGlyLysProGly 1345

Qy     2917 CCAGCCTGGTTCCCCCCACTTCCA--------------------CTCCCCTCTACTCTC 2955
            ||||||    |||    ||||||                         ||||||
Db     1346 ProAlaGlyLeuProGlyLeuProGlyMetLysGlyGluSerGlyLeuProGlyProGln 1365

Qy     2956 TCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCC 3012
                  ||||||          ||||||   :::         |||      |||
Db     1366 GlyProAlaGlyLeu---ProGlyLeuProGlyLeuLysGlyGluProGlyLeuPro 1383
```

RESULT 35
A54849
collagen alpha 1(VII) chain precursor - human
N;Alternate names: procollagen alpha 1(VII) chain
C;Species: Homo sapiens (man)
C;Date: 04-Nov-1994 #sequence_revision 04-Nov-1994 #text_change 20-Sep-1999
C;Accession: A54849; PH0844; S16316; I56328; A30296; I84686

R;Christiano, A.M.; Greenspan, D.S.; Lee, S.; Uitto, J.
J. Biol. Chem. 269, 20256-20262, 1994
A;Title: Cloning of human type VII collagen. Complete primary sequence of the alpha1(VII
A;Reference number: A54849; MUID:94327588; PMID:8051117
A;Accession: A54849
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-2944 <CHR>
A;Cross-references: GB:L02870; NID:g987124; PIDN:AAA75438.1; PID:g987125
R;Tanaka, T.; Takahashi, K.; Furukawa, F.; Imamura, S.
Biochem. Biophys. Res. Commun. 183, 958-963, 1992
A;Title: Molecular cloning and characterization of type VII collagen cDNA.
A;Reference number: PH0844; MUID:92231902; PMID:1567409
A;Accession: PH0844
A;Molecule type: mRNA
A;Residues: 'EFR',340-475,'RALSTASHSTLCWRATRWHPCNRGSHWTRAACEPCNRPASHRAARAG',524-528,'C',
A;Cross-references: DDBJ:D11152; DDBJ:D13694; NID:g453698; PIDN:BAA02853.1; PID:g453699
A;Experimental source: keratinocyte
A;Note: the authors translated the codon ACC for residues 394 and 397 as Tyr
R;Parente, M.G.; Chung, L.C.; Ryynaenen, J.; Woodley, D.T.; Wynn, K.C.; Bauer, E.A.; Mat
Proc. Natl. Acad. Sci. U.S.A. 88, 6931-6935, 1991
A;Title: Human type VII collagen: cDNA cloning and chromosomal mapping of the gene.
A;Reference number: S16316; MUID:91334380; PMID:1871109
A;Accession: S16316
A;Molecule type: mRNA
A;Residues: 815-892,'E',894-1439 <PAR>
A;Cross-references: GB:M65158; GB:S49017; NID:g180914; PIDN:AAA96439.1; PID:g180915
A;Experimental source: keratinocyte
R;Gammon, W.R.; Abernethy, M.L.; Padilla, K.M.; Prisayanh, P.S.; Cook, M.E.; Wright, J.;
J. Invest. Dermatol. 99, 691-696, 1992
A;Title: Noncollagenous (NC1) domain of collagen VII resembles multidomain adhesion prot
A;Reference number: I56328; MUID:93107742; PMID:1469284
A;Accession: I56328
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 'EFR',372-517,'DV',520-540,'W',542-1255 <RES>
A;Cross-references: GB:S51236; NID:g262308; PIDN:AAB24637.1; PID:g262309
R;Seltzer, J.L.; Eisen, A.Z.; Bauer, E.A.; Morris, N.P.; Glanville, R.W.; Burgeson, R.E.
J. Biol. Chem. 264, 3822-3826, 1989
A;Title: Cleavage of type VII collagen by interstitial collagenase and type IV collagena
A;Reference number: A30296; MUID:89139437; PMID:2537292
A;Accession: A30296
A;Molecule type: protein
A;Residues: 'A',1240-1246,'G',1248-1250,'XE',1253-1255,'Q',1257,'E';2032,'C',2034-2041;'
A;Note: two reported peptides cannot be reliably located
R;Greenspan, D.S.
Hum. Mol. Genet. 2, 273-278, 1993
A;Title: The carboxyl-terminal half of type VII collagen, including the non-collagenous
A;Reference number: I48103; MUID:93271985; PMID:8499916
A;Accession: I84686
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 2395-2871,'S',2873-2944 <RE2>
A;Cross-references: GB:L06862; NID:g388713; PIDN:AAA89196.1; PID:g388714
R;Christiano, A.M.; Ryynaenen, M.; Uitto, J.
Proc. Natl. Acad. Sci. U.S.A. 91, 3549-3553, 1994
A;Title: Dominant dystrophic epidermolysis bullosa: identification of a Gly --> Ser subs
A;Reference number: A55255; MUID:94224777; PMID:8170945
A;Contents: annotation
C;Comment: Prolines and lysines at the third position of the tripeptide repeating unit (
ed and subsequently O-glycosylated.
C;Genetics:
A;Gene: GDB:COL7A1; EBR1; EBD1; EB
A;Cross-references: GDB:128750; OMIM:120120
A;Map position: 3p21.3-3p21.3
A;Note: defects in this gene can result in dominant and recessive dystrophic epidermolys
A;Note: there are 118 introns
C;Complex: type VII collagen is probably a homotrimer
C;Function:
A;Description: structural component of extracellular polymer associated with anchoring f
C;Superfamily: unassigned collagens; animal Kunitz-type proteinase inhibitor homology; f
C;Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxyproli
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-2944/Product: collagen alpha 1(VII) chain #status predicted <MAT>
F;17-1253/Domain: amino-terminal nonhelical #status predicted <NC1>
F;36-201/Domain: von Willebrand factor type A repeat homology <VWA1>
F;231-318/Domain: fibronectin type III repeat homology <FN1>
F;327-413/Domain: fibronectin type III repeat homology <FN2>
F;414-502/Domain: fibronectin type III repeat homology <FN3>
F;508-593/Domain: fibronectin type III repeat homology <FN4>
F;598-683/Domain: fibronectin type III repeat homology <FN5>
F;686-771/Domain: fibronectin type III repeat homology <FN6>
F;776-862/Domain: fibronectin type III repeat homology <FN7>
F;864-952/Domain: fibronectin type III repeat homology <FN8>
F;954-1045/Domain: fibronectin type III repeat homology <FN9>
F;1052-1219/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1170-1172/Region: cell attachment (R-G-D) motif
F;1189-1253/Region: cysteine/proline-rich
F;1254-2783/Region: interrupted helical
F;1334-1336/Region: cell attachment (R-G-D) motif
F;2008-2010/Region: cell attachment (R-G-D) motif
F;2553-2555/Region: cell attachment (R-G-D) motif
F;2784-2944/Domain: carboxyl-terminal nonhelical #status predicted <NC2>
F;2876-2929/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>
F;337,786,1109/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;2167,2176,2185,2188,2664,2667,2673/Modified site: 4-hydroxyproline (Pro) #status ex
F;2625,2631/Modified site: 5-hydroxylysine (Lys) #status experimental
F;2625,2631/Binding site: carbohydrate (Lys) (covalent) #status experimental
F;2634,2802,2804/Disulfide bonds: interchain #status predicted

```
Alignment Scores:
Pred. No.:              2.5e-11       Length:        2944
Score:                  312.50        Matches:       260
Percent Similarity:     31.26%        Conservative:  37
Best Local Similarity:  27.37%        Mismatches:    362
Query Match:            5.02%         Indels:        291
DB:                     2             Gaps:          53

US-09-759-143-110 (1-3410) x A54849 (1-2944)

Qy     2489 CCTGGTTCTTGTGTGTTG-----------CCCCTCAGGACTCTTCCCCTACAAATAAGT 2442
             ||||||     ||||||          ||||||||
Db     1158 ProGlyValMetValLeuLeuValAspGluProLeuArg-------------------- 1170

Qy     2441 CATATGTTCAAATCCCATGGAGGAGTGTTTCATCCTAGAAACTCCCATGCAAGAGCTACA 2382
                               |||   :::|||   |||
Db     1171 -----------------GlyAspIlePheSerPro------------------------ 1176

Qy     2381 TTAAACGAAGCTGCAGGTTAAGGGG---------------------------------- 2357
            :::   |||||| |||   |||
Db     1177 IleArgGluAla-GlnAlaSerGlyLeuAsnValValMetLeuGlyMetAlaGlyAlaAs 1196

Qy     2356 ---------CTTAGAGATGGGAAACCAGGT---GACTGAGTT---------TATTCAGC 2319
                     ||||||      ||||||   |||   |||          ::::::
Db     1196 pProGluGlnLeuArgArgLeuAlaProGlyMetAspSerValGlnThrPhePheAlaVa 1216

Qy     2318 TCCCAAAAACCCTTCTCTA------GGTGTGTCTCAACTAGGAGGCTAGCTGTTAACCCT 2265
                     |||||||||          ||||||   |||          |||
Db     1216 lAspAspGlyProSerLeuAspGlnAlaValSerGlyLeuAlaThrAlaLeuCysGlnAl 1236

Qy     2264 GAGCCTGGGTAATCCACCTGCAGAGTCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCT 2205
             |||          |||          |||          :::   |||   |||
Db     1236 aSerPheThrThrGlnProArgProGluPro---CysProValTyrCysProLysGly-- 1254

Qy     2204 CCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGA 2145
                                                    |||   |||   |||   ::
Db     1255 ---------------------------------------GlnLysGlyGluProGlyGl 1261

Qy     2144 CTGGGGAGAGAGGAGA------GGGACGCCCCAGCCCCCAGCTGTGCAGCTACGCACCTC 2091
            :   |||   |||          |||   |||      |||   :::   ||||||
Db     1261 uMetGlyLeuArgGlyGlnValGlyProProGlyAspProGlyLeuProGlyArgThrGl 1281

Qy     2090 AGCAGCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCC 2031
             |||        |||   :::   |||          ||||||            |||:
Db     1281 yAlaPro------GlyProGlnGlyProPro------GlySer-AlaThrAlaLysGlyG 1297
```

```
Qy   2030 AGCCCGGCAGCCCCATGGGGCTAACAGGAGCGG---GGAGCTGGGACCCAGTGAGGCAGG 1974
          ::   |||  |||  |||      ||| |||  |||:::      :::  |||
Db   1297 luArgGlyPhePro---GlyAlaAspGly-ArgProGlySerProGlyArgAlaGlyAsn 1315

Qy   1973 CCCTCCACCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGT 1914
           |||   ||||||
Db   1316 ProGlyThrProGlyAla---------------------------------------- 1321

Qy   1913 CAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCCTGCGGCAGACACCA 1854
                                                 |||||||||:::  :::|||
Db   1322 -------------------------------------ProGlyLeuLysGlySerPro 1328

Qy   1853 TATAGGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCATAAACAGGGATGGGGCCA 1794
                         |||    |||:::     |||          |||
Db   1329 GlyLeuProGlyProArgGlyAspProGlyGluArgGlyProArgGlyProLysGlyGlu 1348

Qy   1793 CCTGGGACAGCAGGAAGG-----------CACTATCCAGGATGGCGAGGTCCAGGCAGA 1746
           ||||||   |||:::                 ||||||   |||
Db   1349 ProGlyAlaProGlyGlnValIleGlyGlyGluGlyProGlyLeuProGlyArgLysGly 1368

Qy   1745 TGCCCCGGCCCGGAA---CCACCCTGGCCTCGGTGGGCTCACCCACCACCACACGTACGG 1689
            |||||||||     ||||||  |||||| |||                 ||
Db   1369 AspProGlyProSerGlyProProGlyProArg-Gly--------------ProLeuGl 1383

Qy   1688 AGACATCACAGGCAGAGGCCCC-------------------------GCAGAGCGCGG 1656
           ||||      |||||||||                          :::::||
Db   1383 yAspProGlyProArgGlyProProGlyLeuProGlyThrAlaMetLysGlyAspLysGl 1403

Qy   1655 GTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGGGAAGGGAGCTC 1596
          |   |||  |||  |||       ||||||  |||                   |||
Db   1403 yAspArg-GlyGluArgGly---ProProGlyPro-----------------GlyGluG 1416

Qy   1595 CAGGCTTAGGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCTAGCACCTCCAG 1536
            |||:::   ||||||
Db   1416 lyGlyIleAlaProGly---------------------------------------- 1421

Qy   1535 TGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGTAGAGGGAGGCCAGTG 1476
               ||| |||:::       |||  :  |||   |||||||||   :::|||  ||||||
Db   1422 --GluPro-GlyLeu---ProGlyLeuProGlySerProGlyProGlnGly---ProVal 1438

Qy   1475 TGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAGCTGTCA 1416
                |||      |||   :::      :::     |||        |||
Db   1439 GlyProProGlyLysLysGlyGluLysGlyAspSerGluAspGlyAlaProGlyLeuPro 1458

Qy   1415 CCACGGCCACACTGTGGGACAGGCATGTGGCACCGGCAGCCACAGGGAAAGCTG------ 1362
            :::|||                           |||    |||    |||    :::
Db   1459 GlyGlnProGlySerProGlyGluGlnGlyProArgGlyProProGlyAlaIleGlyPro 1478

Qy   1361 ---------CCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCTGCACCAGCCGGT 1311
                     :::|||     |||   |||              |||
Db   1479 LysGlyAspArgGlyPheProGlyProLeuGlyGluAlaGlyGluLysGlyGluArgGly 1498

Qy   1310 CCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACAGCCCCAGGCTGCCCA 1251
           |||                |||||||         ||||||:::   |||   |||
Db   1499 Pro---------------ProGly---------ProAlaGlySerArg---GlyLeuPro 1509

Qy   1250 TCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGC 1191
                                   ||||||       |||             |||
Db   1510 GlyValAlaGlyArgProGlyAlaLysGlyProGluGlyProProGlyProThrGlyArg 1529

Qy   1190 CCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCA 1131
               |||          |||     |||      |||              |||
Db   1530 GlnGlyGluLysGlyGluProGlyArgProGlyAspProAla------------ValVal 1545

Qy   1130 TCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGC------ 1077
                      |||   |||:::                   |||    |||
Db   1546 GlyProAlaValAlaGlyProLysGlyGluLysGlyAspValGlyProAlaGlyProArg 1565

Qy   1076 AGCACAGCTGGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGGCCC 1017
                |||     ||||||   |||   ||||||          |||     |||
Db   1566 GlyAlaThrGlyValGlnGlyGluArgGlyProProGlyLeuValLeuProGlyAspPro 1585

Qy   1016 GGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGGTGG 957
          |||                           |||||||||         |||     |||
Db   1586 GlyProLysGlyAspProGlyAsp---ArgGlyProIleGly---LeuThrGlyArgAla 1603

Qy    956 GGC-----------------------------------CCAGCGCTGCCTCCTCAG 936
          |||                                    |||      |||
Db   1604 GlyProProGlyAspSerGlyProProGlyGluLysGlyAspProGlyArgProGlyPro 1623

Qy    935 CCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGC 876
          |||   |||               |||   |||      |||       |||    |||
Db   1624 ProGlyProValGlyProArgGlyArgAspGlyGluValGlyGluLysGlyAspGluGly 1643

Qy    875 ACTCCT------CCTGGGTGCCCAGGTAGG-----------------------GGGCCA 846
              |||     ||||||  |||||||:::                          |||
Db   1644 ProProGlyAspProGlyLeuProGlyLysAlaGlyGluArgGlyLeuArgGlyAlaPro 1663

Qy    845 GGGCACTGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGACTGA 786
          |||   |||                          |||    |||       |||
Db   1664 GlyValThrGlyProVal---------------GlyGluLysGlyAspGlnGlyAspPro 1678

Qy    785 TCATGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGTCCGGGTCCCGGAAGAGGTCAG 726
                    |||       ||||||  |||||||||         ||||||:::
Db   1679 GlyGluAspGlyArgAsnGlySerProGlySerSerGlyProLysGlyAspArgGlyGlu 1698

Qy    725 AGAGCA------GGGCCTCCA----------------------------GTGGAGTGA 702
                      |||||||||                              |||
Db   1699 ProGlyProProGlyProProGlyArgLeuValAspThrGlyProGlyAlaArgGluLys 1718

Qy    701 AGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGC---------------CCAGGATGA 657
             ||||||   |||          ||||||                    ||||||
Db   1719 GlyGluProGlyAspArgGlyGlnGluGlyProArgGlyProLysGlyAspProGly--- 1737

Qy    656 GCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGGCCC 597
                   |||||||||      ||||||
Db   1738 ---------LeuProGlyAla------ProGly--------------------------- 1743

Qy    596 TTGGGATGAGAAAGAGGCTCAGCAGGATGC---------------CCAAGGACAGTGCCC 552
                   |||||||||      |||                  |||:::     |||
Db   1744 ---------GluArgGlyIleGluGlyPheArgGlyProProGlyProGlnGlyAspPro 1760

Qy    551 AGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTAGGA 492
                |||      |||
Db   1761 GlyValArgGlyProAlaGlyGluLys--------------------------------- 1769

Qy    491 GCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCATGGTCATGAACTTCT 432
             |||   |||      |||   ||||||                          |||
Db   1770 ---GlyAspArgGlyProProGlyLeuAspGly--------------------ArgSer 1781

Qy    431 CCTCTACCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCCAAACACA 372
            |||           |||   |||                         |||:::
Db   1782 GlyLeuAspGlyLysProGlyAlaAlaGly--------------------ProSerGly 1794

Qy    371 CCTCCAGGCCAAAGG---TTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGT---GCC 318
          |||   |||              ||||||   |||   |||   ||||||  |||
Db   1795 ProAsnGlyAlaAlaGlyLysAlaGlyAspProGlyArgAspGlyLeuProGlyLeuArg 1814

Qy    317 GCAGCAGGCGGCTCACCCACAGCCTCTGGACCATAGTGGG-------------------- 278
                   |||   |||      |||||||||   |||
Db   1815 GlyGluGlnGlyLeuProGlyProSerGlyPro-ProGlyLeuProGlyLysProGlyGl 1834

Qy    277 ----------CCAGGCGGGTAGGGCTCAGGGGGCCGTTCAGGCACTCCA---GAACTGCT 231
                    ||||||    |||     |||       |||   |||   |||
Db   1834 uAspGlyLysProGlyLeuAsnGlyLysAsnGlyGluProGlyAspProGlyGluAspGl 1854

Qy    230 TCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAACTGCCTAG 171
           |||   |||            ::::::  ||||||              |||    |||::
Db   1854 yArgLysGlyGluLysGlyAspSerGlyAlaSerGlyArgGluGlyArgAspGlyProLy 1874

Qy    170 GAATCAGCCAGGCGCC---CATTTCTGCCAGCCCTTTGGTGCCGGTCCAGCTTCTCAGCC 114
          :   :::  ||||||               |||   |||  |||          |||     |
Db   1874 sGlyGluArgGlyAlaProGlyIleLeuGlyProGlnGly-ProProGlyLeuProGlyP 1894

Qy    113 CATGCTCAACACCTGCT---GCTGTGGGGCACCTCAGTGGGGACACGTCTCATCACTCAG 57
           ||         |||                     :::   |||   |||        |
```

```
Db       1894 roValGlyProProGlyGlnGlyPheProGlyValProGlyGlyThrGlyProLysGlyA 1914

Qy         56 ATCCTGGC-------------------------------------CGAGGCGCGCGGC 36
              ||   |||                                     ||||||   |||
Db       1914 spArgGlyGluThrGlySerLysGlyGluGlnGlyLeuProGlyGluArgGlyLeuArgG 1934

Qy         35 TGTCACCCGGAGCC 22
                   ||||||:::
Db       1934 lyGluProGlySer 1938
```

RESULT 36
CGHU2V
collagen alpha 2(V) chain precursor - human
C;Species: Homo sapiens (man)
C;Date: 31-Jul-1989 #sequence_revision 28-Jul-1995 #text_change 31-Dec-2000
C;Accession: A31427; A54555; S43643; A25874; I55239; I59025; A25374; A30017
R;Woodbury, D.; Benson-Chanda, V.; Ramirez, F.
J. Biol. Chem. 264, 2735-2738, 1989
A;Title: Amino-terminal propeptide of human pro-alpha2(V) collagen conforms to the struc
A;Reference number: A31427; MUID:89123368; PMID:2914927
A;Accession: A31427
A;Molecule type: mRNA
A;Residues: 1-463 <WOO>
A;Cross-references: GB:J04478; NID:g179697; PIDN:AAA51859.1; PID:g179698
A;Experimental source: placenta
R;Greenspan, D.S.; Lee, S.T.; Lee, B.S.; Hoffman, G.G.
Gene Expr. 1, 29-39, 1991
A;Title: Homology between alpha2(V) and alpha1(III) collagen promoters and evidence for
A;Reference number: A54555; MUID:92314691; PMID:1820205
A;Accession: A54555
A;Molecule type: DNA
A;Residues: 1-32 <GRE>
A;Cross-references: GB:M58529; NID:g180834; PIDN:AAC41699.1; PID:g553235
R;Moradi-Ameli, M.; Rousseau, J.C.; Kleman, J.P.; Champliaud, M.F.; Boutillon, M.M.; Ber
Eur. J. Biochem. 221, 987-995, 1994
A;Title: Diversity in the processing events at the N-terminus of type-V collagen.
A;Reference number: S43642; MUID:94237164; PMID:8181482
A;Accession: S43643
A;Molecule type: protein
A;Residues: 288-291,'P',293-294,'X',296-297;606,'X',608-617 <MOR>
R;Weil, D.; Bernard, M.; Gargano, S.; Ramirez, F.
Nucleic Acids Res. 15, 181-198, 1987
A;Title: The pro alpha 2(V) collagen gene is evolutionarily related to the major fibrill
A;Reference number: A25874; MUID:87146331; PMID:3029669
A;Accession: A25874
A;Molecule type: mRNA; DNA
A;Residues: 398-1496 <WEI>
A;Cross-references: GB:X04758; NID:g29588; PIDN:CAA28454.1; PID:g1340175
A;Experimental source: rhabdomyosarcoma cell line
R;Myers, J.C.; Loidl, H.R.; Stolle, C.A.; Seyer, J.M.
J. Biol. Chem. 260, 5533-5541, 1985
A;Title: Partial covalent structure of the human alpha 2 type V collagen chain.
A;Reference number: I55239; MUID:85182703; PMID:2985598
A;Accession: I55239
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1002-1226 <RE2>
A;Cross-references: GB:M10956; NID:g180427; PIDN:AAA52007.1; PID:g180428
A;Note: part of this sequence were determined by protein sequencing
R;Emanuel, B.S.; Cannizzaro, L.A.; Seyer, J.M.; Myers, J.C.
Proc. Natl. Acad. Sci. U.S.A. 82, 3385-3389, 1985
A;Title: Human alpha 1(III) and alpha 2(V) procollagen genes are located on the long arm
A;Reference number: I59025; MUID:85216505; PMID:3858826
A;Accession: I59025
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1003-1034 <RES>
A;Cross-references: GB:M11135; NID:g179693; PIDN:AAA51857.1; PID:g179694
A;Note: part of this sequence were determined by protein sequencing
R;Myers, J.C.; Loidl, H.R.; Seyer, J.M.; Dion, A.S.
J. Biol. Chem. 260, 11216-11222, 1985
A;Title: Complete primary structure of the human alpha-2 type V procollagen COOH-termina
A;Reference number: A25374; MUID:85289337; PMID:2411731
A;Accession: A25374
A;Molecule type: mRNA
A;Residues: 1227-1417,'T',1419-1437,'S',1439-1496 <MYE>
A;Cross-references: GB:M11718; NID:g180912; PIDN:AAA52058.1; PID:g180913
A;Experimental source: normal fibroblasts
R;Tsipouras, P.; Schwartz, R.C.; Liddell, A.C.; Salkeld, C.S.; Weil, D.; Ramirez, F.
Genomics 3, 275-277, 1988
A;Title: Genetic distance of two fibrillar collagen loci, COL3A1 and COL5A2, located
A;Reference number: A30017; MUID:89138450; PMID:3224983
A;Accession: A30017
A;Molecule type: DNA
A;Residues: 1449-1463,'E',1465-1495,'A' <TSI>
A;Cross-references: GB:J03051; NID:g179695; PIDN:AAA51858.1; PID:g179696
A;Note: the authors translated the codon GAA for residue 1460 as Gln, and GAG for res
C;Comment: Prolines and lysines at the third position of the tripeptide repeating uni
are 5-hydroxylated and subsequently O-glycosylated.
C;Comment: The amino-terminal propeptide domain appears not to be completely cleaved.
C;Genetics:
A;Gene: GDB:COL5A2
A;Cross-references: GDB:119064; OMIM:120190
A;Map position: 2q31-2q31
A;Introns: 33/1; 812/3; 830/3; 848/3; 884/3; 902/3; 922/3; 974/3; 1046/3; 1064/3; 144
C;Complex: type V collagen may be a homotrimer of alpha 1(V) chains (see PIR:CGHU1V),
alpha 2(V) chain and one alpha 3(V) chain, initially linked by disulfide bonds among
ength, is formed with desmosine cross-links made from lysine and allysine residues
C;Function:
A;Description: structural component of extracellular fibrous polymer associated with
A;Note: may play a role in controlling the lateral growth of collagen I fibrils
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo
C;Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxypr
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-1250/Product: collagen alpha 2(V) chain #status predicted <MAT>
F;27-193/Domain: amino-terminal propeptide (uncleaved) #status predicted <NPP>
F;27-108/Region: nonhelical
F;40-99/Domain: von Willebrand factor type C repeat homology <VWC>
F;109-186/Region: helical
F;187-208/Region: nonhelical
F;209-1225/Region: helical
F;503-505/Region: cell attachment (R-G-D) motif
F;941-943/Region: cell attachment (R-G-D) motif
F;1064-1066/Region: cell attachment (R-G-D) motif
F;1067-1069/Region: cell attachment (R-G-D) motif
F;1097-1099/Region: cell attachment (R-G-D) motif
F;1124-1126/Region: cell attachment (R-G-D) motif
F;1133-1135/Region: cell attachment (R-G-D) motif
F;1225-1250/Region: carboxyl-terminal nonhelical telopeptide
F;1251-1496/Domain: carboxyl-terminal propeptide #status predicted <CPP>
F;1269-1496/Domain: fibrillar collagen carboxyl-terminal homology <FCC>
F;27/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic
F;193-194/Cleavage site: Ala-Gln (procollagen N-endopeptidase) #status predicted
F;194/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predi
F;201/Modified site: allysine (Lys) #status predicted
F;290,293,296,608,614,1004,1007,1013,1028,1034/Modified site: 4-hydroxyproline (Pro)
F;299,1139/Modified site: 5-hydroxylysine (Lys) #status predicted
F;299,1139/Binding site: carbohydrate (Lys) (covalent) #status predicted
F;1025/Modified site: 5-hydroxylysine (Lys) #status experimental
F;1250-1251/Cleavage site: Glu-Asp (procollagen C-endopeptidase) #status predicted
F;1259,1397/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;1293,1299,1325/Disulfide bonds: interchain #status predicted
F;1333-1494,1402-1447/Disulfide bonds: #status predicted

Alignment Scores:

Pred. No.:	2.68e-11	Length:	1496
Score:	312.00	Matches:	226
Percent Similarity:	30.80%	Conservative:	37
Best Local Similarity:	26.46%	Mismatches:	255
Query Match:	4.86%	Indels:	336
DB:	1	Gaps:	51

US-09-759-143-110 (1-3410) x CGHU2V (1-1496)

```
Qy         26 CCGGGTGAC-----------------------AGCCGCGCGCCTCGG------CCAGGA 55
              |||||||||                       ||||||   |||         ||||||
```

```
Db   242 ProGlyAspProGlyProMetGlyProIleGlySerArgGlyProGluGlyProProGly 261

Qy    56 TCTGAGTGATGAGACGTGTCCCCACTGAGG------------------------------ 85
                       |||       |||    |||
Db   262 LysProGlyGluAspGlyGluProGlyArgAsnGlyAsnProGlyGluValGlyPheAla 281

Qy    86 ---------------------TGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAGC 124
                              |||    |||    |||:::       |||||||:::
Db   282 GlySerProGlyAlaArgGlyPheProGlyAlaProGlyLeuPro---GlyLeuLysGly 300

Qy   125 TGGACCGGCACCAAAGGGCTG--------------------------------------- 145
            |||    |||||||||
Db   301 HisArgGlyHisLysGlyLeuGluGlyProLysGlyGluValGlyAlaProGlySerLys 320

Qy   146 ------------------GCAGAAATGGGCGCCTGG------CTGATTCCTAGGCAGTTG 181
                           |||||||||         |||    ||||||    :::
Db   321 GlyGluAlaGlyProThrGlyProMetGlyAlaMetGlyProLeuGlyProArgGlyMet 340

Qy   182 GCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGG 241
                  ||| |||                ||||||      :::|||   :::   ||
Db   341 ProGlyGluArg-GlyArgLeuGlyProGlnGlyAlaProGlyGlnArgGlyAlaHisGl 360

Qy   242 AGTGCCTGAACGGCCC-------------------------------------------- 257
         |:::|||   :::|||
Db   360 yMetProGlyLysProGlyProMetGlyProLeuGlyIleProGlySerSerGlyPhePr 380

Qy   257 ------------------------------------------------------------ 257

Db   380 oGlyAsnProGlyMetLysGlyGluAlaGlyProThrGlyAlaArgGlyProGluGlyPr 400

Qy   258 -------------CCTGAGCCCTACCCGCCTGGCCCACTATGGT------------CCAG 292
                        |||      |||||||||||||::: |||            |||
Db   400 oGlnGlyGlnArgGlyGluThrGlyProProGlyProVal-GlySerProGlyLeuProG 420

Qy   293 AGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAA------------AGCCCAGCTCTTG 340
          |||:::|||                   |||:::            ||||||     :
Db   420 lyAlaIleGlyThrAspGlyThrProGlyProLysGlyProThrGlySerProGlyThrS 440

Qy   341 CTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTG 400
         ::|||  |||    :::        |||    |||        |||    :::
Db   440 erGlyProProGlySerAlaGlyProProGlySerProGlyProGln------------- 455

Qy   401 CCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATT 460
              |||          |||       |||         |||         |||
Db   456 --GlySerThrGlyProGlnGlyAsnSerGlyLeuProGlyAsp------ProGlyPheL 473

Qy   461 GGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCG- 519
                  |||||||||          |||   |||          |||   |||:::
Db   473 ysGlyGluAlaGlyProLysGlyGluProGlyProHisGlyIleGlnGlyProIleGlyP 493

Qy   520 ----------------------------------------TGGACGCTATGGCCGCCG- 537
                                                  ||||||    ||||||
Db   493 roProGlyGluGluGlyLysArgGlyProArgGlyAspProGlyThrLeuGlyProProG 513

Qy   538 --CCGGCCCTTCATCTGGGCACTGTC------CTTGGGCATCCT---GCTGAGCCTCTTT 586
           |||         |||            |||   |||    ::::::     |||
Db   513 lyProValGlyGluArgGlyAlaProGlyAsnArgGlyPheProGlySerAspGlyLeuP 533

Qy   587 CTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAG 646
               |||||||||            |||      |||    ||||||      |||
Db   533 roGlyProLysGlyAlaGlnGlyGluArgGlyProVal---GlySerSerGlyProLysG 552

Qy   647 CTGGCACTGCTCATCCTGGGCGT---GGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
                        |||||||||   |||      ||||||       ||||||
Db   552 lySerGlnGlyAspProGlyArgProGlyGluProGlyLeu------ProGlyAlaArgG 570

Qy   704 ACTCCACTGGAGGCCCTGCT--------------------CTCTGACCTCTTCCGGGAC 742
               ||||||   |||                      |||    ||||||     |
Db   570 lyLeuThrGlyAsnProGlyValGlnGlyProGluGlyLysLeuGlyProLeuGlyAlaP 590

Qy   743 CCGGA-----------CCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGT 790
          |||||          |||     ||||||                            |
Db   590 roGlyGluAspGlyArgProGlyProProGly---------------------------S 601

Qy   791 CTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCC 850
         ||   |||:::    |||    |||           :::|||          ||||||
Db   601 erIleGlyIleLysGlyGlnProGlyThr---------MetGly------LeuProGlyP 616

Qy   851 CCCTA--------------CCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACC 895
          ||                 ||||||   ||||||                    :::|
Db   616 roLysGlySerAsnGlyAspProGlyLysProGlyGluAla------4-----GlyAsnP 632

Qy   896 CTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGC 955
         ||   :::|||       |||                |||    |||      |||      ||||
Db   632 roGlyValProGlyGlnArgGly---------AlaProGlyLysAspGlyLysValGlyP 649

Qy   956 CCCAC------------------------CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCC 991
         ||:::                         |||      |||       ||||||
Db   649 roTyrGlyProProGlyProProGlyLeuArgGlyGluArgGlyGluGlnGlyPro---- 667

Qy   992 TTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTG 1051
                              |||||||||   ||||||       |||    |||
Db   668 ----------------------ProGlyProThrGlyPhe------GlnGlyHisProG 678

Qy  1052 CTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTG 1111
              |||     |||
Db   678 lyProProGlyProProGly---------------------------------------- 684

Qy  1112 GCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTG 1171
                         :::|||                            |||      |
Db   685 -----------------GluGly---------------------------GlyLysProG 690

Qy  1172 GGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACAC 1231
          ||          |||||||||         :::        |||      ||||||
Db   690 lyAspGlnGlyValProGlyGlyProGlyAla---ValGlyProLeuGlyPro------- 706

Qy  1232 TATGATGAAGGCGTTCGGATGGGCAGCCTGGG------GCTGTTCCTGCAGTGCGCCATC 1285
                   |||      |||   ||||||           |||    :::      |
Db   707 --------ArgGlyGluArgGlyAsnProGlyGluArgGlyGluProGlyIleThrGlyL 724

Qy  1286 TCCCT---------GGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACT 1336
          |||||         |||:::       |||||||||||   |||
Db   724 euProGlyGluLysGlyMetAlaGlyGlyHisGlyProAspGly---------------- 738

Qy  1337 CGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTG 1396
                                          |||    |||              |||
Db   739 --------------------------------ProLysGly-----------SerProG 744

Qy  1397 TCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCC 1456
            |||          ||||||:::
Db   744 lyProSerGlyThrProGlyAspThr---------------------------------- 752

Qy  1457 CTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTG 1516
              |||    |||     |||:::|||         |||    ||||||    |||
Db   753 --GlyProProGlyLeuGln---GlyMetProGlyGluArgGlyIleAlaGlyThrProG 771

Qy  1517 CCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGC--- 1573
                  :::||||||    ||||||          :::    :::        |||:::
Db   771 ly-Pro---LysGlyAspArgGlyGlyIleGlyGluLysGlyAlaGluGlyThrAlaGly 789

Qy  1574 ---------------TTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACAC 1618
                        ||||||||||||      |||    |||    |||    |||
Db   790 AsnAspGlyAlaGlyGlyLeuProGlyProLeu---GlyProProGlyProAlaGlyLeu 808

Qy  1619 GTGGGTGCTGGAGGCAGT------------GGCCTGCTCCCACCTCCA------------ 1654
         :::|||     |||                |||||||:::    ||||||
Db   809 LeuGlyGluLysGlyGluProGlyProArgGlyLeuValGlyProProGlySerArgGly 828

Qy  1655 ---CCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTG------------ 1699
            |||       |||    :::              :::|||
Db   829 AsnProGlySerArgGlyGluAsnGlyProThrGlyAlaValGlyPheAlaGlyProGln 848

Qy  1700 --------------------GTGGTGGGTGAGCCCACCGAGGCCAGGTGGTTCCGGGC 1738
                             |||    |||||||||    |||  |||  |||       |||
Db   849 GlySerAspGlyGlnProGlyValLysGlyGluProGlyGlu-Pro-GlyGlnLysGlyA 868
```

```
Qy       1739 CGGGGCATCTGCCTGGACCTCGC-----------CATCCTGGATAGTGCCTTCCTGCTG 1786
                           |||||||||:::             ||||||
Db        868 spAlaGlySerProGlyProGlnGlyLeuAlaGlySerProGly--------------- 882

Qy       1787 TCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCAT-----------TGTCCAGCTCAGC 1834
                |||  ||||||     |||  ||||||                    |
Db        883 --ProHisGlyProAsnGlyValProGlyLeuLysGlyGlyArgGlyThrGlnGlyProP 902

Qy       1835 CAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTT 1894
              ||               |||         |||||||||
Db        902 ro-----------------GlyAlaThrGlyPheProGlySerAla------------ 911

Qy       1895 GCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCC 1954

Db        911 ------------------------------------------------------------ 911

Qy       1955 AGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCC 2014
                            |||     |||  |||          ||| |||  |||      |||
Db        912 -------GlyArgValGlyProProGlyProAlaGlyAla-ProGlyProAlaGlyPro 928

Qy       2015 ATGGGGCTGCCGGGC------TGGCCGCCA 2038
              :::|||  ||||||        ||||||
Db        929 LeuGlyGluProGlyLysGluGlyProPro 938
```

RESULT 37
I49607
procollagen type V alpha 2 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 13-Aug-1999
C;Accession: I49607
R;Andrikopoulos, K.; Suzuki, H.R.; Solursh, M.; Ramirez, F.
Dev. Dyn. 195, 113-120, 1992
A;Title: Localization of pro-alpha 2(V) collagen transcripts in the tissues of the devel
A;Reference number: I49607; MUID:93214071; PMID:1297453
A;Accession: I49607
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1497 <RES>
A;Cross-references: GB:L02918; NID:g309180; PIDN:AAA37440.1; PID:g309181
C;Genetics:
A;Gene: Col5a-2
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
F;39-98/Domain: von Willebrand factor type C repeat homology <VWC>
F;1270-1497/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

```
Alignment Scores:
Pred. No.:              2.88e-11       Length:          1497
Score:                  311.50         Matches:         225
Percent Similarity:     30.02%         Conservative:    32
Best Local Similarity:  26.29%         Mismatches:      258
Query Match:            4.85%          Indels:          341
DB:                     2              Gaps:            47
```

US-09-759-143-110 (1-3410) x I49607 (1-1497)

```
Qy         20 CTGGCTCCGGGTGACAGCCGCGCGCCTCGG------CCAGGATCTGAGTGATGAGACGTG 73
              :::  |||    ||||||  |||          ||||||         |||
Db        246 MetGlyProIleGlySerArgGlyProGluGlyProProGlyLysProGlyGluAspGly 265

Qy         74 TCCCCACTGAGG------------------------------------------------ 85
                |||   |||
Db        266 GluProGlyArgAsnGlyAsnThrGlyGluValGlyPheSerGlySerProGlyAlaArg 285

Qy         86 ---TGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAGCTGGACCGGCACCAAAGGG 142
                  |||   |||   |||:::      ||||||:::          |||   ||||||
Db        286 GlyPheProGlyAlaProGlyLeuPro---GlyLeuLysGlyHisArgGlyHisLysGly 304

Qy        143 CTG--------------------------------------------------------- 145
              |||
Db        305 LeuGluGlyProLysGlyGluIleGlyAlaProGlyAlaLysGlyGluAlaGlyProThr 324

Qy        146 GCAGAAATGGGCGCCTGG------CTGATTCCTAGGCAGTTGGCGGCAGCAAGGAGGAGA 199
                 |||||||||          |||   ||||||  :::          ||| |||
Db        325 GlyProMetGlyAlaMetGlyProLeuGlyProArgGlyMetProGlyGluArg-GlyAr 344

Qy        200 GGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAGTGCCTGAACGGCCC-- 257
                       ||||||     :::|||   :::   |||:::|||   :::|||
Db        344 gLeuGlyProGlnGlyAlaProGlyLysArgGlyAlaHisGlyMetProGlyLysProGl 364

Qy        257 ------------------------------------------------------------ 257

Db        364 yProMetGlyProLeuGlyIleProGlySerSerGlyPheProGlyAsnProGlyMetLy 384

Qy        258 -------------------------------------------------------CCTGA 262
                                                                          ||
Db        384 sGlyGluArgGlyProHisGlyAlaArgGlyProGluGlyProGlnGlyGlnArgGlyGl 404

Qy        263 GCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGT----------------------- 299
              |     ||||||||||||     |||:::|||
Db        404 uThrGlyProProGlyProAlaGlySerGlnGlyLeuProGlyAlaValGlyThrAspGl 424

Qy        300 ---------------GGGTGAGCCGCCTGCTGCGGCACC---GGAAAGCCCAGCTCTTG 340
                             |||          ||||||  |||    |||   ||||
Db        424 yThrProGlyArgLysGlyAlaThrGlySerAlaGlyThrSerGly-ProProGlyLeuA 444

Qy        341 CTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTG 400
               |||||  |||   :::    |||  |||
Db        444 laGlyProProGlySerProGlyProGlnGly-------------------------- 454

Qy        401 CCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCT---GGGC 457
                   |||        |||  ||||||:::           |||    |||    |||
Db        455 -----SerThrGlyProGlnGlyIleArgGlyGlnSerGlyAspProGlyValProGlyP 473

Qy        458 ATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGG 517
                         |||||||||        |||   |||          |||   |||:::
Db        473 heLysGlyGluAlaGlyProLysGlyGluProGlyProHisGlyIleGlnGlyProIleG 493

Qy        518 CGTGGACGCTATGGCCGCCG-------------------CCGGCCCTTCATCTGGGCA 556
                          ||||||                    |||              ||||
Db        493 ly-----------ProProGlyGluGluGlyLysArgGlyProArgGlyAspProGlyT 509

Qy        557 CTGTC--------------CTTGGGCATCCTGCTGAGCCT------------------ 582
              |||||              :::|||        |||
Db        509 hrValGlyProProGlyProMetGlyGluArgGlyAlaProGlyAsnArgGlyPheProG 529

Qy        583 ----------CTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGAT 631
                        |||    |||||||||        |||      |||    ||||
Db        529 lySerAspGlyLeuProGlyProLysGlyAlaGlnGlyGluArgGlyProVal---GlyS 548

Qy        632 CCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGT---GGGGCTGCTGGACTTCTGT 688
              ||     |||       |||     |||||||||    |||      ||||||
Db        548 erSerGlyProLysGlyGlyGlnGlyAspProGlyArgProGlyGluProGlyLeu---- 566

Qy        689 GGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT-------------------CTCT 727
                 ||||||          ||||||   |||                          |||
Db        567 --ProGlyAlaArgGlyLeuThrGlyAsnProGlyValGlnGlyProGluGlyLysLeuG 586

Qy        728 GACCTCTTCCGGGACCCGGA-----------CCACTGTCGCCAGGCCTACTCTGTCTAT 775
                 ||||||      ||||||              |||      ||||||   :::  :::
Db        586 lyProLeuGlyAlaProGlyGluAspGlyArgProGlyProProGlySerIleGlyIleA 606

Qy        776 GCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGAC 835
                        |||   |||:::||||||          :::      ||||||
Db        606 rgGlyGlnProGlySerMetGlyValProGlyProLysGlySerSerGlyAspLeuGlyL 626

Qy        836 ACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGA------GGAGTGCCTCTTTGGCCTG 889
                                ||||||    |||         |||||||||
Db        626 ys-----------------ProGlyGluAlaGlyAsnAlaGlyValPro---------- 636

Qy        890 CTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCG 949
                                                     |||   |||      |||
Db        637 ------------------------GlyGlnArgGlyAlaProGlyLysAspGlyGluV 648

Qy        950 CTGGGCCCCAC------------------------CGAGCCAGCAGAAGGGCTGTCGGCC 985
                 ||||||                                   |||       ||||
Db        648 alGlyProSerGlyProValGlyProProGlyLeuAlaGlyGluArgGlyGluAlaGlyP 668
```

```
Qy     986 CCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGC 1045
           ||                                ||||||||   ||||||        |||
Db     668 ro--------------------------ProGlyProThrGlyPhe------GlnGlyL 677

Qy    1046 GCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTC 1105
             |||      |||      |||
Db     677 euProGlyProProGlyProProGly---------------------------------- 685

Qy    1106 TTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGAT 1165
                                 :::|||                           |||
Db     686 -----------------------GluGly---------------------------GlyL 689

Qy    1166 TTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGG 1225
                 |||       |||||||||          :::        |||      ||||||
Db     689 ysAlaGlyAspGlnGlyValProGlyGluProGlyAla---ValGlyProLeuGlyPro- 707

Qy    1226 AGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGG------GCTGTTCCTGCAGTGC 1279
                            |||        |||   ||||||         |||    :::
Db     708 -------------ArgGlyGluArgGlyAsnProGlyGluArgGlyGluProGlyIleT 723

Qy    1280 GCCATCTCCCT---------GGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTC 1330
                 ||||||          |||:::       |||||||||||     |||
Db     723 hrGlyLeuProGlyGluLysGlyMetAlaGlyGlyHisGlyProAspGly---------- 739

Qy    1331 GGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACA 1390
                                                   |||   |||
Db     740 ------------------------------------ProLysGly-----------A 743

Qy    1391 TGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTC 1450
              |||    |||         ||||||:::
Db     743 snProGlyProThrGlyThrIleGlyAspThr---------------------------- 753

Qy    1451 TCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTG 1510
                          |||   |||    |||:::|||         |||   ||||||
Db     754 --------GlyProProGlyLeuGln---GlyMetProGlyGluArgGlyIleAlaGlyT 770

Qy    1511 TTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACC 1570
              |||           :::||||||   ||||||        :::   :::      |||
Db     770 hrProGly-Pro---LysGlyAspArgGlyGlyIleGlyGluLysGlyAlaGluGlyThr 788

Qy    1571 AGC-----------------TTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAAT 1612
           :::                 ||||||||||||       |||   |||   |||
Db     789 AlaGlyAsnAspGlyAlaArgGlyLeuProGlyProLeu---GlyProProGlyProAla 807

Qy    1613 GGACACGTGGGTGCTGGAGGCAGT-----------GGCCTGCTCCCACCTCCA------ 1654
           |||   :::||||||   |||              |||||||:::   ||||||
Db     808 GlyLeuLeuGlyAlaProGlyGluProGlyProArgGlyLeuValGlyProProGlySer 827

Qy    1655 ---------CCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTG------ 1699
                    |||         |||    :::                :::|||
Db     828 ArgGlyAsnProGlySerArgGlyGluAsnGlyProThrGlyAlaValGlyPheAlaGly 847

Qy    1700 -------------------------GTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTT 1732
                                    |||    |||||||||    |||  |||  |||
Db     848 ProGlnGlySerAspGlyGlnProGlyValLysGlyGluProGlyGlu-Pro-GlyGlnL 867

Qy    1733 CCGGGCCGGGGCATCTGCCTGGACCTCGC-----------CATCCTGGATAGTGCCTTC 1780
             |||           |||||||||:::                 ||||||
Db     867 ysGlyAspAlaGlySerProGlyProGlnGlyLeuAlaGlySerProGly---------- 883

Qy    1781 CTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCAT-----------TGTCCAG 1828
                      |||   ||||||      |||    ||||||
Db     884 --------ProHisGlyProHisGlyValProGlyLeuLysGlyGlyArgGlyThrGlnG 901

Qy    1829 CTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATT 1888
              |||                    |||          |||||||||
Db     901 lyProPro-----------------GlyAlaThrGlyPheProGlySerAla------- 912

Qy    1889 TACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAA 1948

Db     912 ------------------------------------------------------------ 912

Qy    1949 ACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTT 2008
                         |||      |||   |||                 |||  |||   |||
Db     913 --------------GlyArgValGlyProProGlyProAlaGlyAla-ProGlyProAla 927

Qy    2009 AGCCCCATGGGGCTGCCGGGC------TGGCCGCCA 2038
              |||   |||   ||||||       ||||||
Db     928 GlyProAlaGlyGluProGlyLysGluGlyProPro 939
```

RESULT 38
A45748
collagen alpha 1(VII) chain - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
C;Accession: A45748
R;Li, K.; Christiano, A.M.; Copeland, N.G.; Gilbert, D.J.; Chu, M.L.; Jenkins, N.A.;
Genomics 16, 733-739, 1993
A;Title: cDNA cloning and chromosomal mapping of the mouse type VII collagen gene (Co
A;Reference number: A45748; MUID:93315168; PMID:8325648
A;Accession: A45748
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-920 <LI1>
A;Cross-references: GB:S63654; NID:g386656; PIDN:AAB27492.1; PID:g386657
A;Experimental source: epidermal keratinocyte
A;Note: sequence extracted from NCBI backbone (NCBIN:135000, NCBIP:135001)
C;Superfamily: unassigned collagens

```
Alignment Scores:
Pred. No.:               3.08e-11     Length:          920
Score:                   311.00       Matches:         208
Percent Similarity:      30.18%       Conservative:    31
Best Local Similarity:   26.26%       Mismatches:      319
Query Match:             5.00%        Indels:          235
DB:                      2            Gaps:            34

US-09-759-143-110 (1-3410) x A45748 (1-920)

Qy    2368 CAGGTTAAGGGGCTTAGAGATGGG---AAACCAGGTGACTGAGTTTATTCAGCTCCCAAA 2312
            |||::::::   |||      |||      |||                 ::::::
Db      62 GlnLeuArgArgLeuAlaProGlyThrAspProIleGlnAsnPhePheAlaValAspAsn 81

Qy    2311 AACCCTTCTCTA------GGTGTGTCTCAACTAGGAGGCTAGCTGTTAACCCTGAGCCTG 2258
              |||   |||            ||||||   |||          |||       ::::::
Db      82 GlyProGlyLeuAspArgAlaValSerAspLeuAlaValAlaLeuCysGlnAlaAlaVal 101

Qy    2257 GGTAATCCACCTGCAGAGTCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGTA 2198
                    |||          |||      :::      |||   |||
Db     102 ThrIleGluProGlnThrGlyPro---CysAlaValHisCysProLysGly--------- 117

Qy    2197 TAAGTCCAGACTGAAACCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGA 2138
                 :::   |||        :::   |||||||||
Db     118 ---GlnLysGlyGluProGlyValThrGlyLeuGlnGly-------------------- 129

Qy    2137 GAGAGGAGAGGGACGCCCCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGT 2078
              :::   |||   |||   ||||||   :::         ||||||   |||
Db     130 ---GlnAlaGlyProProGlyProProGlyLeuProGlyArgThrGlyAlaPro------ 146

Qy    2077 GGCAGCAGAGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCC 2018
                                                            |||      |||
Db     147 ------------------------------------------GlyProGlnGlyProPro 152

Qy    2017 CATGGGGCTAACAGGAGCGGGGAGCTGGGA---CCCAGTGAGGCAGGCCCTCCA------ 1967
                             ||||||   |||   |||         |||||||||
Db     153 GlySerThrGlnAlaLysGlyGluArgGlyPheProGlyProGluGlyProProGlySer 172

Qy    1966 ---------CCCCAATGTGCTGGAAGTTT-------------TCTACGCTGAGTATTTGG 1929
                       |||      ||||||                ||||||
Db     173 ProGlyLeuProGlyValProGlySerProGlyIleLysGlySerThr------------ 188

Qy    1928 CCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCA--- 1872
                                                         ||||||   |||
Db     189 ------------------------------------------GlyArgProGlyProArg 194
```

```
Qy       1871 GGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCA 1812
              |||   :::                     |||         |||      |||:::
Db        195 GlyGluGlnGlyGluArgGlyProGlnGlyProLysGlyGluProGlyGluProGlyGln 214

Qy       1811 TAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCAC--------------------- 1773
                 ||||||  |||  ||||||      |||:::
Db        215 IleThrGlyGlyGlyGlyProGlyPheProGlyLysLysGlyAspProGlyProSerGly 234

Qy       1772 TATCCAGGATGGCGAGGTCCAGGCAGATGCCCCGGCCCGGAA---CCACCCTGGCCTCGG 1716
                 ||||||  |||||||||         |||||||||     ||| |||||||||
Db        235 ProProGlySerArgGlyProValGlyAspProGlyProArgGlyPro-ProGlyLeuPr 254

Qy       1715 TGGGCTCACCCACCACCACACGTACGGAGACATCACAGGCAGAGGCCCCGCAGAGCGCGG 1656
                |||:::::                 ||||||      |||||| |||           |
Db        254 oGlyIleSerValLysGlyAspLysGlyAspArgGlyGluArgGly-ProProGlyProG 274

Qy       1655 GTGGAGGTGGGAGCAGG--------------------------CCACTGCCTCCAGCAC 1623
              ||  |||  |||                               |||     |||   |
Db        274 lyIleGlyAlaSerGluGlnGlyAspProGlyLeuProGlyLeuProGlySerProGlyP 294

Qy       1622 CCACGTGTCCATTAGGGAAG------------------------------------GGAG 1599
              ||    |||  |||:::                                      |||
Db        294 roGlnGlyProAlaGlyArgProGlyGluLysGlyGluLysGlyAspCysGluAspGlyG 314

Qy       1598 CTCCAGGCTTA-----------GGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCAC 1551
                |||||||||            ||||||     :::       :::|||    :::
Db        314 lyProGlyLeuProGlyGlnProGlyProProGlyGluProGlyLeuArgGlyAlaProG 334

Qy       1550 TGCTAGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGT 1491
                :::       |||          |||   |||    |||||||||      ||||||
Db        334 lyMetThrGlyProLysGlyAspArgGlyLeu-ThrGlyThrProGlyGluProGlyVal 353

Qy       1490 AGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGG 1431
              :::|||
Db        354 LysGlyGlu--------------------------------------------------- 356

Qy       1430 CGGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACCGGCAGCCACAG 1371
                                 |||  |||   |||
Db        357 -----------------ArgGlyHisProGly---------------------------- 361

Qy       1370 GGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCTGCACCAGCCGGT 1311
                                                                |||   |||
Db        362 ------------------------------------------------ProValGly 364

Qy       1310 CCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACAGCCCCAGGCTGCCCA 1251
              |||                                                |||   |||
Db        365 Pro---------------------------------------------GlnGlyLeuPro 369

Qy       1250 TCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGC 1191
                          |||      |||  ||||||        |||              |||
Db        370 GlyAlaAlaGlyHisProGlyValGluGlyProGluGlyProProGlyProThrGlyArg 389

Qy       1190 CCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCA 1131
                |||          |||         |||         |||               |||
Db        390 ArgGlyGluLysGlyGluProGlyArgProGlyAspProAlaValGly---------Pro 406

Qy       1130 TCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACA 1071
                          |||:::       :::       ||||||
Db        407 GlyGlyAlaGlyAlaLysGlyGluLysGlyGluAlaGlyLeuProGlyProArgGlyAla 426

Qy       1070 GCTGGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGGCCCGGCATG 1011
              :::|||:::   ||||||    |||   ||||||          |||      ||||||
Db        427 SerGlySerLysGlyGluGlnGlyAlaProGlyLeuAlaLeuProGlyAspProGlyPro 446

Qy       1010 GACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGC--- 954
                                 |||||||||          |||       |||   |||
Db        447 LysGlyAspProGlyAsp---ArgGlyProIleGly---LeuThrGlyArgAlaGlyPro 464

Qy        953 ------------------------------------CCAGCGCTGCCTCCTCAGCCACCA 930
                                                  |||       |||       |||
Db        465 ThrGlyAspSerGlyProProGlyGluLysGlyGluProGlyArgProGlySerProGly 484

Qy        929 GCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGCACTCCT 870
                 |||         |||   |||       |||      |||      |||   |||
Db        485 ProValGlyProArgGlyArgAspGlyGluAlaGlyGluLysGlyAspGluGlyIlePro 504

Qy        869 ------CCTGGGTGCCCAGGTAGG-----------------------GGGCCAGGGCAC 840
                    ||||||   ||||||:::                         ||||||
Db        505 GlyGluProGlyLeuProGlyLysAlaGlyGluArgGlyLeuArgGlyAlaProGlyPro 524

Qy        839 TGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGACTGATCATGA 780
                    |||                     |||      |||       |||
Db        525 ArgGlyProVal---------------GlyGluLysGlyAspGlnGlyAspProGlyGlu 539

Qy        779 AGGCATAGACAGAGTAGGCCTGGCGACAGTGGTCCGGGTCCCGGAAGAGGTCAGAGAGCA 720
                    |||         ||||||   |||||||||         ||||||:::
Db        540 AspGlyArgAsnGlySerProGlySerSerGlyProLysGlyAspArgGlyGluProGly 559

Qy        719 ------GGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGC 666
                    |||||||||                   |||      |||
Db        560 ProProGlyProProGlyArgLeuValAspAlaGlyIleGluSerArgAspLysGlyGlu 579

Qy        665 CCAGGATGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTAGCC 606
                ||||||       |||                |||||||||          :::
Db        580 ProGlyGlnGluGlyProArgGlyProLysGlyAspProGlyProProGlyValSerGly 599

Qy        605 AGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCAAGGACAGTGCCCAGATGA 546
                 |||   :::            ||||||      |||   |||:::      |||
Db        600 GluArgGlyIleAspGlyLeuArgGlyProProGly---ProGlnGlyAspProGlyVal 618

Qy        545 AGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTAGGAGCGGGA 486
                 |||         |||
Db        619 ArgGlyProAlaGlyAspLys-----------------------------------Gly 626

Qy        485 CACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCATGGTCATGAACTTCTCCTCTA 426
                 |||      |||   ||||||                          |||   |||
Db        627 AspArgGlyProProGlyLeuAspGly-------------------ArgSerGlyLeu 639

Qy        425 CCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCCAAACACACCTCCA 366
                       |||   |||                                   ||||||
Db        640 AspGlyLysProGlyAlaProGly----------------------------ProPro 649

Qy        365 GGC-----------CAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGTGCC 318
              |||                      ||||||   |||   |||   ||||||   |||
Db        650 GlyLeuHisGlyAlaSerGlyLysAlaGlyAspProGlyArgAspGlyLeuProGlyLeu 669

Qy        317 GCAGCAGGCGGCTCACCCACAGCCTCTGGACCATAGTGGGCCAGGCGGG---TAGGGCTC 261
                       |||          ||||||    ||||||    |||       |||
Db        670 ArgGlyGluHisGlyPro------ProGlyPro-ProGlyProProGlyValProGlyLy 687

Qy        260 AGGGGGCCGTTCAGGCACTCCAGAACTG 233
                  |||       |||   |||   |||
Db        687 sAlaGlyAspAspGlyLysProGlyLeu 696
```

RESULT 39
CGRT1S
collagen alpha 1(I) chain - rat (tentative sequence) (fragments)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 31-Mar-2000
C;Accession: A90559; A90552; A92029; A90353; A90566; A90357; A90362; A90379; A91209;
R;Bornstein, P.
Biochemistry 8, 63-71, 1969
A;Title: Comparative sequence studies of rat skin and tendon collagen. II. The absenc
A;Reference number: A90559; MUID:69155173; PMID:5777344
A;Contents: CNBr0 and CNBr1
A;Accession: A90559
A;Molecule type: protein
A;Residues: 1-19 <BO1>
A;Experimental source: tendon
A;Note: sequences from skin and tendon appear to be identical
A;Note: the amino-terminal tetrapeptide may be removed by limited proteolysis during
R;Kang, A.H.; Bornstein, P.; Piez, K.A.
Biochemistry 6, 788-795, 1967
A;Title: The amino acid sequence of peptides from the cross-linking region of rat ski
A;Reference number: A90552; MUID:67162268; PMID:5337886
A;Contents: CNBr1

A;Accession: A90552
A;Molecule type: protein
A;Residues: 5-19 <KAN>
A;Experimental source: skin
R;Bornstein, P.
J. Biol. Chem. 242, 2572-2574, 1967
A;Title: The incomplete hydroxylation of individual prolyl residues in collagen.
A;Reference number: A92029; MUID:67165368; PMID:4290711
A;Contents: CNBr2
A;Accession: A92029
A;Molecule type: protein
A;Residues: 20-55 <BO2>
A;Experimental source: skin and tendon
R;Butler, W.T.; Ponds, S.L.
Biochemistry 10, 2076-2081, 1971
A;Title: Chemical studies on the cyanogen bromide peptides of rat skin collagen. Amino a
A;Reference number: A90353; MUID:71263178; PMID:4327399
A;Contents: CNBr4
A;Accession: A90353
A;Molecule type: protein
A;Residues: 56-102 <BU1>
A;Experimental source: skin
R;Butler, W.T.
Biochemistry 9, 44-50, 1970
A;Title: Chemical studies on the cyanogen bromide peptides of rat skin collagen. The cov
A;Reference number: A90566; MUID:70085124; PMID:5411206
A;Contents: CNBr5
A;Accession: A90566
A;Molecule type: protein
A;Residues: 103-139 <BU2>
A;Experimental source: skin
R;Balian, G.; Click, E.M.; Bornstein, P.
Biochemistry 10, 4470-4478, 1971
A;Title: Structure of rat skin collagen alpha1-CB8. Amino acid sequence of the hydroxyla
A;Reference number: A90357; MUID:72136131; PMID:4335087
A;Contents: CNBr8
A;Accession: A90357
A;Molecule type: protein
A;Residues: 140-238 <BA1>
A;Experimental source: skin
R;Balian, G.; Click, E.M.; Hermodson, M.A.; Bornstein, P.
Biochemistry 11, 3798-3806, 1972
A;Title: Structure of rat skin collagen alpha1-CB8. Amino acid sequence of the hydroxyla
A;Reference number: A90362; MUID:73006942; PMID:4342027
A;Contents: CNBr8
A;Accession: A90362
A;Molecule type: protein
A;Residues: 239-418 <BA2>
A;Experimental source: skin
R;Butler, W.T.; Underwood, S.P.; Finch Jr., J.E.
Biochemistry 13, 2946-2953, 1974
A;Title: Chemical studies on the cyanogen bromide peptides of rat skin collagen. Amino a
A;Reference number: A90379; MUID:74271984; PMID:4366532
A;Contents: CNBr3
A;Accession: A90379
A;Molecule type: protein
A;Residues: 419-567 <BU3>
A;Experimental source: skin
R;Stoltz, M.; Timpl, R.; Furthmayr, H.; Kuehn, K.
Eur. J. Biochem. 37, 287-294, 1973
A;Title: Structural and immunogenic properties of a major antigenic determinant in neutr
A;Reference number: A91209; MUID:74011954; PMID:4126850
A;Contents: CNBr6
A;Accession: A91209
A;Molecule type: protein
A;Residues: 568-651 <ST1>
A;Experimental source: skin
A;Note: this region probably corresponds to positions 949-1032 of the alpha 1(I) chain
A;Note: the major antigenic determinant (of neutral salt-extracted rat skin collagen) in
R;Stoltz, M.; Timpl, R.; Kuehn, K.
FEBS Lett. 26, 61-65, 1972
A;Title: Non-helical regions in rat collagen alpha1-chain.
A;Reference number: A91385; MUID:73049495; PMID:4636751
A;Contents: CNBr6
A;Accession: A91385
A;Molecule type: protein
A;Residues: 651-671 <ST2>
A;Experimental source: skin
A;Note: the composition of peptides comprising residues 1-9 and 1-19 confirms the seq
A;Note: this region (residues 651-671 above) probably corresponds to positions 1032-1
C;Comment: Prolines and lysines at the third position of the tripeptide repeating uni
ed and subsequently O-glycosylated.
C;Comment: The order of the nine CNBr peptides in the alpha 1(I) chain of rat skin co
C;Comment: The complete chain contains 1052 residues.
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo
C;Keywords: blocked amino end; coiled coil; extracellular matrix; glycoprotein; hydro
F;1/Modified site: blocked amino end (Glx) (probably pyrrolidone carboxylic acid) #st
F;9/Modified site: allysine (Lys) #status experimental
F;103,424,547/Binding site: carbohydrate (Lys) (covalent) #status experimental
F;103/Modified site: 5-hydroxylysine (Lys) #status experimental
F;424,547/Modified site: 5-hydroxylysine (Lys) (partial) #status experimental

```
Alignment Scores:
Pred. No.:                3.55e-11    Length:         671
Score:                    310.00      Matches:        213
Percent Similarity:       33.33%      Conservative:   35
Best Local Similarity:    28.63%      Mismatches:     279
Query Match:              4.98%       Indels:         218
DB:                       1           Gaps:           40

US-09-759-143-110 (1-3410) x CGRT1S (1-671)

Qy     2131 AGAGGGACGCCCCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGC 2072
             ||||||   |||  ||||||                           ||||||
Db       25 ArgGlyLeuProGlyProProGlyAlaPro---GlyProGlnGlyPheGlnGlyProPro 43

Qy     2071 AGAGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGG 2012
             ||||||          |||::::::           |||           |||
Db       44 GlyGluProGlyGluProGlyAlaSerGlyProMetGlyProArgGlyProProGlyPro 63

Qy     2011 GCTAACAGGAGCGGGGAGCTGGGA--------------------CCCAGTGAGGCAGGC 1973
                 ::::::|||:::   |||                   |||   :::   |||
Db       64 ProGlyLysAsnGlyAspAspGlyGluAlaGlyLysProGlyArgProGlyGlnArgGly 83

Qy     1972 CCTCCA-----------------------CCCCAATGTGCTGGAAG------------- 1950
            ||||||                       |||    ||||||
Db       84 ProProGlyProGlnGlyAlaArgGlyLeuProGlyThrAlaGlyLeuProGlyMetLys 103

Qy     1949 -----------TTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCT 1902
                        ||||||  |||
Db      104 GlyHisArgGlyPheSerGlyLeuAsp--------------------------------- 112

Qy     1901 GTGTAGCAAAGTAAATGGCGACCAGACCCA---GGCCTGCGGCAGACACCATATAGGCAG 1845
                      :::|||          |||   |||   :::       |||
Db      113 ------GlyAlaLysGlyAsnThrGlyProAlaGlyProLysGlyGluProGlySerPro 130

Qy     1844 TGACAGACTGGCTGAGCTGGACAATGGAGCCCATAAACAGGGATG-------------- 1800
              :::   |||     ||||||    |||      |||:::
Db      131 GlyGluAsnGlyAlaProGlyGlnMetGlyPro---ArgGlyLeuProGlyGluArgGly 149

Qy     1799 ------GGGCCACCTGGGACAGCAGGAAGGCACTATCCAGGATGGCGAGGTCCAGGCAGA 1746
                  ||||||||||||:::||||||             ||||||   |||   |||
Db      150 ArgProGlyProProGlySerAlaGlyAlaArgGly-AspAspGlyAlaValGlyAlaAl 169

Qy     1745 TGCCCCGGCCCGGAACCACCCTGGCCTCGGTGGGCTCACCCACCACCACGTACGGAGA 1686
              |||  ||||||  |||              |||   |||
Db      169 aGlyPro-ProGlyProThr---GlyProThrGlyProProGlyPheProGlyAlaAlaG 188

Qy     1685 CATCACAGGCAGAGGCC---CCGCAGAGCGCGGGTGGAGGT-----------GGGAGCA 1641
               ::::::   ||||||   ||||||   |||   |||               |||   |
Db      188 lyAlaLysGlyGluAlaGlyProGlnGlyAlaArgGlySerGluGlyProGlnGlyValA 208

Qy     1640 GG------CCACTGCCTCCAGCACCCACGTGTCCATTAGGG---AAGGGAGCTCCAGGCT 1590
             ||      |||   ||||||   |||                |||    |||   ||||||
Db      208 rgGlyGluProGlyProProGlyProAlaGlyAlaAlaGlyProAlaGlyAsnProGlyA 228
```

```
Qy     1589 TAGGG------CCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCTAGCACCTCCAG 1536
                       ||||||  |||   ::::::|||  |||
Db      228 laAspGlyGlnProGlyAlaLys-GlyAlaAsnGlyAlaPro------------------ 241

Qy     1535 TGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGTAGAGGGAGGCCAGTG 1476
                     ||||||  ||||||  |||   ||||||  ||||||
Db      242 ---------GlyIle---AlaGlyAlaProGlyPheProGlyAlaArgGly--------- 254

Qy     1475 TGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAGCTGTCA 1416
                  :::|||  ||||||  |||  ||||||  |||  |||
Db      255 ---ProSerGlyPro-GlnGlyProSerGlyAlaProGlyProLysGlyAsnSerGlyGl 273

Qy     1415 CCACGGCCACACTGTGGGACAGGCATGTGGCACCGGCAGCCACAGGGAAAGCTGCCACAC 1356
               |||      |||        ||||||:::  :::|||:::
Db      273 uProGly---AlaProGlyAsnLysGlyAspThrGlyAla--LysGlyGlu--------- 288

Qy     1355 TGGCCAAATAGACTGCTCGAGTGCCGAATCGCTGCACCAGCCGGT-----------CCA 1308
                                                ||||||||      |||
Db      289 -----------------------------ProGlyProAlaGlyValGlnGlyProPro 298

Qy     1307 TGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACA---------------GCC 1263
               |||      |||:::      |||
Db      299 GlyProAlaGlyGluGluGlyLysArgGlyAlaArgGlyGluProGlyProSerGlyLeu 318

Qy     1262 CCAGGCTGCCCATCCGAACGC---------------------------CTTCATCATAGT 1230
            ||||||  |||  ||||||
Db      319 ProGlyProProGlyGluArgGlyGlyProGlySerArgGlyPheProGlyAlaAspGly 338

Qy     1229 GTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTACAGCCCCTCGCCCA 1170
            |||:::|||||||:::  ||||||            ||||||  |||
Db      339 ValAlaGlyProLysGlyProAlaGlyGluArgGlySerProGlyProAlaGlyProLys 358

Qy     1169 CGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGCTGCACAGCTCAGCCA 1110
               :::|||
Db      359 GlySerPro--------------------------------------------------- 361

Qy     1109 CGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACAGCTGGTGCAGCCGGGGAAGCA 1050
                                                ||||||  |||||||||
Db      362 --------------------------------GlyGluAlaGlyArgProGlyGluAla 370

Qy     1049 GGGCGCCCAGGT-----------TCCGGAAAGCCAAGCGGGCCCGGCATGGACAGCAGT 1002
            |||  ||||||         :::|||  |||     ||||||  |||
Db      371 GlyLeuProGlyAlaLysGlyLeuThrGlySerProGlySerProGlyProAsp------ 388

Qy     1001 GGGGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCT 942
            |||  |||         |||         |||  |||||||||  |||
Db      389 GlyLysThrGlyProProGlyProAlaGly******GlyArgProGlyProAlaGlyPro 408

Qy      941 CCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAA 882
            |||                              |||  |||  ||||||
Db      409 Pro------------------------------GlyAlaArgGlyGlnAlaGlyVal 417

Qy      881 AGAGGCACTCCT---------------CCTGGGTGCCCAGGTAGGGGGCCAGGGCACTGG 837
               |||  |||              |||  ||||||:::
Db      418 MetGlyPheProGlyProLysGlyThrAlaGlyGluProGlyLysAlaGlyGluArgGly 437

Qy      836 TGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGG 777
               |||         |||  |||     |||         :::
Db      438 ValProGlyProProGlyAlaValGlyProAlaGlyLysAspGlyGlu---------Ala 454

Qy      776 CATAGACAGAGTAGGCCTGGCGACAGTGGTCCGGGTCCCGGAAGAGGTCAGAGA------ 723
               |||   ||||||  :::||||||      ||||||:::::
Db      455 GlyAlaGlnGlyAlaProGlyProAlaGlyProAlaGlyGluArgGlyGluGlnGlyPro 474

Qy      722 ----------------------------------GCAGGGCCTCCAGTGGAGTGAAGC 699
                                              ||||||||||||  |||
Db      475 AlaGlySerProGlyPheGlnGlyLeuProGlyProAlaGlyProProGlyGluAlaGly 494

Qy      698 ACACCTGGCCACAGAAGTCCAGCAGCCCCACGCCCAGGATGAGCAGTGCCAGCTCCA--- 642
            ||||||                                       ||||||
Db      495 LysProGly******--------------------------GlyValProGlyAspLeu 505

Qy      641 GGGGCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGA 582
            ||||||      ||||||   :::                |||          ||||||
Db      506 GlyAla------ProGlyProSerGlyAlaArgGlyGluArgGlyPheProGlyGluArg 523

Qy      581 GGCTCAGCAGGATGCCCAAGGACAGTG---CCCAGATGAAGGGCCGGCGGCCATAGC 525
            |||      |||  |||            ||||||        |||
Db      524 GlyValGlnGlyProProGlyProAlaGlyProArgGlyAsnAsnGlyAlaProGly*** 543

Qy      524 GTCCACGCCAGTGGTCACTGGCTGAGCCTAGGAGCG---GGACACAGACCAGGCCCAGCA 468
               |||  |||     :::  ||||||  |||    :::     |||
Db      544 ***GlyAlaLysGlyAspThrGlyAlaProGlyAlaProGlySerGlnGlyAlaProGly 563

Qy      467 CTGGACCAATGC------------------CCAGCACCATGGTCATGAACTTCTCCTCTA 426
            |||:::                         |||  |||
Db      564 Leu***GlyMetSerGlyLeu***GlyProProGlyProProGlySerProGly****** 583

Qy      425 CCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCCAAACACACCTCCA 366
               |||   |||:::    |||                     |||     ||||||
Db      584 GlyProSerGlyAlaSerGlyProAla---------------GlyProArgGlyProPro 598

Qy      365 GGCCAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGTGCCGCAGCAGGCGGC 306
            |||      ||||||  |||   :::  ||||||  |||
Db      599 GlySer------AlaGlySerProGlyLys***GlyLeu***GlyLeu------------ 612

Qy      305 TCACCCACAGCCTCTGGACCATAGTGGGCCA---GGCGGGTAGGGCTCAGGGGGCCGTTC 249
               |||      ||||||   ||||||  |||      |||     |||  ||
Db      613 ---ProGlyProIleGlyPro-ProGlyProArgGlyArgThrGly***AlaGlyProSe 631

Qy      248 AGGCACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCT 191
            ||||  |||                     |||           ||||||
Db      631 rGlyProPro---------------------GlyProProGlyProProGlyProPro 643
```

RESULT 40
A40333
collagen alpha 1'(II) chain precursor - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 16-Jul-1999
C;Accession: A40333
R;Su, M.W.; Suzuki, H.R.; Bieker, J.J.; Solursh, M.; Ramirez, F.
J. Cell Biol. 115, 565-575, 1991
A;Title: Expression of two nonallelic type II procollagen genes during Xenopus laevis
A;Reference number: A40333; MUID:92011898; PMID:1918153
A;Accession: A40333
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-1492 <SUA>
A;Cross-references: GB:M63596
A;Note: this sequence is presented as substitutions relative to another sequence in a
es they replace; the appropriate interpretation of the sequence figure was reconstruc
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo
C;Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix
F;37-96/Domain: von Willebrand factor type C repeat homology <VWC>
F;1263-1492/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Alignment Scores:

Pred. No.:	3.8e-11	Length:	1492
Score:	309.50	Matches:	236
Percent Similarity:	32.52%	Conservative:	44
Best Local Similarity:	27.41%	Mismatches:	299
Query Match:	4.97%	Indels:	282
DB:	2	Gaps:	50

US-09-759-143-110 (1-3410) x A40333 (1-1492)

```
Qy     2164 CAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCCCCAGCCCCCAGCTGTG 2105
            |||   |||   |||   |||        :::      |||           ||||||
Db      110 GlnLysGlyGluProGlyAspIleLysAspValValGlyProArgGlyProProGlyPro 129

Qy     2104 CAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGCAAC 2045
            |||                             |||   :::  |||           :::
Db      130 Gln--------------GlyProSerGlyGluGlnGlyProArgGlyGluArgGlyAsp 144

Qy     2044 AGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGGA--- 1988
            :::            |||      |||   |||      |||   ||||||   |||
```

```
Db     145 LysGlyGluThrGlyAlaProGlyProArgGly------ArgAspGlyGluProGlyThr 162

Qy    1987 ---------CCCAGTGAGGCAGGCCCT---------------CCACCCCAATGTGCTGGA 1952
                   |||   |||||||||                  ||||||      |||
Db     163 ProGlyAsnProGlyProAlaGlyProProGlyProProGlyProProGlyLeuGlyGly 182

Qy    1951 AGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAA 1892
           :::|||                                              ::::::
Db     183 AsnPhe----------------------------------------------AlaAlaGln 187

Qy    1891 GTAAATGGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCT 1832
           :::   |||         :::     |||        :::|||  |||      |||
Db     188 MetThrGlyGlyPheAspGluLysAlaGlyGlyAlaGlnMetGly-Val-MetGlnGlyP 207

Qy    1831 GAGCTGGACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGG---- 1776
                 |||        |||              |||||||||||   :::|||
Db     207 roMetGlyProMetGlyPro---------ArgGlyProProGlyProSerGlyAlaProG 224

Qy    1775 -------------------------CACTATCCAGGATGGCGAGGTCCAGGCAGATGCC 1742
                                               ||||||     |||    |||
Db     224 lyProGlnGlyPheGlnGlyAsnProGlyGluProGlyGluSerGlyAlaGlyGlyProM 244

Qy    1741 CCGGCCCGGAA---CCACCCTGGCCTCGGTGGGCTCACCCACCACCACGTACGGAGAC 1685
             ||||||         ||||||   |||     :::|||     |||          :::
Db     244 etGlyProArgGlyProProGlyProSer--GlyLysPro-----------GlyAspAs 259

Qy    1684 ATCACAGGCAGAGGCCCCGCAGAGCGCGGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGC 1625
                :::|||         |||     :::|||        |||           ||||||
Db     259 pGlyGluAlaGlyLysProGlyLysSerGlyGluArgGly-----------ProProGl 275

Qy    1624 ACCCACGTGTCCATTAGGG---AAGGGAGCTCCAGGCTTAGGGCCTGGCAGGAAGCTGGT 1568
            |||           |||       |||   |||||||||   ||||||   |||
Db     275 yProGlnGlyAlaArgGlyPheProGlyThrProGlyLeu---ProGlyValLys----- 292

Qy    1567 CATCAGGCTGTCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAACAC 1508
                          |||                                ||| |
Db     293 ----------Gly-His--------------------------------ArgGlyTyrP 298

Qy    1507 CTGCTTCTCCCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAA 1448
            ||      ||||||:::|||               ||||||:::|||||||   :::
Db     298 roGlyLeuAspGlyGlyLysGly------------GluAlaGlyAlaAlaGlyAlaLys- 313

Qy    1447 GGTGAACCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGT 1388
           ||||||  |||      ||| ||||||             :::       |||   ::
Db     314 GlyGluSerGlyAlaSerGly-GluAlaGlyAlaProGlyProMetGlyProArgGlyLe 333

Qy    1387 G--------------GCACCGGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGAC 1343
           :                  |||   :::::::|||  |||              |||
Db     333 uProGlyGluArgGlyArgProGlySerSerGlyAlaAlaGlyAlaArgGly-AsnAsp- 352

Qy    1342 TGCTCGAGTGCCGAATCGCTGCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGAT 1283
                          :::       ||||||||||||                ||||||
Db     353 ----------GlyLeuProGlyProAlaGlyPro---------------ProGlyProV 364

Qy    1282 GGCGCACTGCAGGAACAGCC------CCAGGCTGCCCATCCGAACGCCTTCATCATAGTG 1229
                  ||||||                 ||||||   |||        :::
Db     364 alGlyProAlaGlyAlaProGlyPheProGlyAlaProGlySerLysGlyGluAlaGlyP 384

Qy    1228 TCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTACAGCCCCTCGCCCAC 1169
             :::|||    |||   |||                |||   :::         |||
Db     384 roThrGlyAlaArgGlyProGluGlyAlaGlnGlyProArgGlyGluSerGlyThrProG 404

Qy    1168 GAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGCTGCACAGCTCAGCCAC 1109
             :::|||                        ||||||   :::   |||
Db     404 lySerPro------------------------GlyProSerGlyAlaSerGlyAsnProG 416

Qy    1108 GAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACAGCTGGTGCAGCCGGGGAAGCAG 1049
                 |||          ||||||   ||||||   :::|||||:::|||    ||||
Db     416 lyThrAspGlyIlePro---GlyAlaLysGlySer---SerGlyAlaSerGlyIleAlaG 434

Qy    1048 GGCGCCCAGGTTCCGGAAAGCCAAGCGGGCCCGGCATGGACAGCAGTGGGGCGACA---- 993
            ||  ||||||        |||   ||||||          ||||||||
Db     434 lyAlaProGlyPheProGlyProArgGlyPro---ProGlyProGlnGlyAlaThrGlyP 453
```

```
Qy     992 --------------------------------------------------AGGAGGG 986
                                                                   |
Db     453 roLeuGlyProLysGlyGluThrGlyAspProGlyIleAlaGlyPheLysGlyGluHisG 473

Qy     985 GGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCTCCT------------- 939
            |||||              ||||||  ||||||       |||  |||
Db     473 lyProLysGly--GluIleGlySerAlaGlyProGlnGlyAla-ProGlyProAlaGlyG 492

Qy     938 ------------------------CAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAG 905
                                   :::|||  |||         |||       |
Db     492 luGluGlyLysArgGlyAlaArgGlyGluProGlyAlaAlaGlyProLeuGlyProProG 512

Qy     904 GAAGATGAGGGTGAGCAGGCCAAAGAGGC-----------------------ACTCCTC 869
            ||   |||   |||  ||||||                                  |||
Db     512 lyGluArgGlyAlaProGlyAsnArgGlyPheProGlyGlnAspGlyLeuAlaGlyProL 532

Qy     868 CTGGGTGCCCAGGT---AGGGGG---CCAGGGCACTGGTGTCCCAGTCAATGGCAGGCAG 815
              |||   ||||||    ||||||   ||||||          |||
Db     532 ysGlyAlaProGlyGluArgGlyValProGlyLeuGlyGlyProLysGlyGlyAsnGlyA 552

Qy     814 GA---GGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTG 758
              |||  ||||||  |||                |||             |||||||
Db     552 spProGlyArgProGlyGluProGlyLeuProGlyAlaArgGlyLeuThrGlyArgProG 572

Qy     757 GCGACAGTGGTCCG-------------GGTCCC-------------GGAAGAGGTCAGAGAG 722
            |||||:::||||||          ||||||          |||          |||
Db     572 lyAspAlaGlyProGlnGlyLysValGlyProSerGlyAlaAlaGlyGluAspGlyArgP 592

Qy     721 CAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGC---- 666
            |||||||||                            |||    :    |||
Db     592 roGlyProPro-------------------------GlyProGlnGlyAlaArgGlyG 603

Qy     665 -------------------CCAGGATGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGG 626
                              ||||||      |||               :::|||
Db     603 lnProGlyValMetGlyPheProGly---------ProLysGlyAlaAsnGlyGluProG 620

Qy     625 GCACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCC 566
            ||  |||                              |||:::|||      |||   |
Db     620 lyLysAlaGly--------------------------GluLysGlyLeuGlyGlyAlaP 631

Qy     565 CA------AGGACAGTGCCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTG 512
            ||      |||   :::|||                |||             |||  |
Db     631 roGlyLeuArgGlyLeuPro------------GlyLysAspGlyGluThrGlyAlaGlnG 647

Qy     511 GTCACTGGCTGAGCCTAGGA---GCGGGACACAGA---------------CCAGGCCCAG 470
            ||               |||   ||||||   |||                  |||||||||:
Db     647 lyProAsnGlyProAlaGlyProAlaGlyGluArgGlyGluGlnGlyProProGlyProS 667

Qy     469 CA---CTGGACCAATGCCCAGCACCATGGTCATGAACTTCTCCTCTACCCCCACTTCCAG 413
            ::                 |||   |||
Db     667 erGlyPheGlnGlyLeuProGlyPro-------------------------------- 675

Qy     412 CAGCAGAGGCGGCACATAGGTGATGCCTGCGGCCAAACACACCT---------------C 368
                                          |||    :::|||                |
Db     676 -------------------------------ProGlySerProGlyGluGlyGlyLysP 685

Qy     367 CAGGCCAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGTGCCGCAGCA---- 312
            |||||    :::         |||   |||        ||||||    :::|||
Db     685 roGlyAspGln------GlyValProGlyGluAlaGlyAlaAlaGlyLeuValGlyProA 703

Qy     311 --------GGCGGCTCACCCACAGCCTCTGGACCATAGTGGGCCAGGCGGG---TAGGGC 263
                    |||     |||              |||   |||||||||    |||      |||
Db     703 rgGlyGluArgGlyPheProGlyGluArgGly-SerSerGlyProGlnGlyLeuGlnGly 722

Qy     262 TCAGGGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCTCGGC------------------ 221
            |||    |||    |||||||||              |||
Db     723 SerArgGlyLeuProGlyThrPro---------------GlyThrAspGlyProLysGly 737

Qy     220 ------------TCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAACTGCCT 173
                          |||   |||          |||        ||||||        :::    :::|||
Db     738 AlaThrGlyProSerGlyProAsnGlyAlaGlnGlyProProGlyLeuGlnGlyMetPro 757
```

```
Qy        172 AGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGT-------------------- 134
                  :::    |||              |||   |||
Db        758 GlyGluArgGlyAlaAlaGlyIleSerGlyProLysGlyAspArgGlyAspThrGlyGlu 777

Qy        133 --------------------------------------------GCCGGTCCAGCTTCT 119
                                                          ||||||
Db        778 LysGlyProGluGlyAlaProGlyLysAspGlySerArgGlyLeuThrGlyProLeuGly 797

Qy        118 CAGCCCATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTCATCACTC 59
               |||           ||||||        |||:::::
Db        798 ProPro------Gly-ProAlaGlyProAsnGlyGluLysGlyGluSerGlyProSerGl 815

Qy         58 AGATCCTGGC---CGAGGCGCGCGGCTGTCACCCGGA 25
                 ||||||    |||||||||   :::||||||
Db        815 yProProGlyIleValGlyAlaArgGlyAlaProGly 827
```

RESULT 41
T14339
sucrose-proton transport protein - carrot
N;Alternate names: sucrose/H+ symporter protein
C;Species: Daucus carota (carrot)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C;Accession: T14339
R;Shakya, R.; Sturm, A.
Plant Physiol. 118, 1473-1480, 1998
A;Title: Characterization of source- and sink-specific sucrose/H+ symporters from carrot
A;Reference number: Z17991; MUID:99063785; PMID:9847123
A;Accession: T14339
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-501 <SHA>
A;Cross-references: EMBL:Y16766; NID:g2969886; PIDN:CAA76367.1; PID:g2969887
A;Experimental source: cultivar Namtaise: leaf
C;Genetics:
A;Gene: SUT1a
C;Superfamily: common tobacco sucrose transport protein

Alignment Scores:

Pred. No.:	4.08e-11	Length:	501
Score:	309.00	Matches:	131
Percent Similarity:	37.44%	Conservative:	94
Best Local Similarity:	21.80%	Mismatches:	219
Query Match:	4.81%	Indels:	157
DB:	2	Gaps:	16

US-09-759-143-110 (1-3410) x T14339 (1-501)

```
Qy        286 GGTCCAGAG--------------------------GCTGTGGGTGAGCCGCCTGCTGCG 318
               |||||||||                         |||      ||||||   ||
Db          3 GlyProGluAlaAspArgAsnArgHisArgGlyGlyAlaThrAlaAlaProProPro-Ar 22

Qy        319 GCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTG------CTAACCTTTGGCCTGGAGGT 372
              |  |||   :::       ||||||:::   :::       :::     |||::::::
Db         22 gSerArgValSerLeuArgLeuLeuLeuArgValAlaSerValAlaCysGlyIleGlnPh 42

Qy        373 GTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGA 432
                       |||      :::::   :::   |||   :::   |||:::|||:::
Db         42 eGlyTrpAlaLeuGlnLeuSerLeuLeuThrProTyrValGlnGluLeuGlyIleProHi 62

Qy        433 GAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCT 492
                 :::   :::::::::         ||||||:::   ||||||:::         |||::
Db         62 sAlaTrpSerSerIleIleTrpLeuCysGlyProLeuSerGlyLeuLeuValGlnProIl 82

Qy        493 CCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTG 552
              :::::|||       ||||||         :::|||||||||||||||||||||||||
Db         82 eValGlyHisMetSerAspGlnCysThrSerLysTyrGlyArgArgArgProPheIleVa 102

Qy        553 GGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGC 612
                 |||         |||:::|||:::::   :::|||      :::      :::
Db        102 lAlaGlyGlyThrAlaIleIleLeuAlaValIleIleIleAlaHisSerAlaAspIleGl 122

Qy        613 AGGGCTGCTG------TGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGG 666
                |||||||||         |||   :::   :::   :::         ::::::||

Db        122 yGlyLeuLeuGlyAspThrAlaAspAsnLysThrMetAlaIleValAlaPheValIleGl 142

Qy        667 CGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTC 726
              |      :::||||||              :::         |||      |||||||||::
Db        142 yPheTrpIleLeuAspValAlaAsnAsnMetThrGlnGlyProCysArgAlaLeuLeuAl 162

Qy        727 TGACCTCTTC---CGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCAT 783
              :||||||          |||            |||   |||   :::   ::::::
Db        162 aAspLeuThrGlyAsnAspAlaArgArgThrArgValAlaAsnAlaTyrPheSerLeuPh 182

Qy        784 GATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGC 843
                :::::::::|||     |||||||||                              ||
Db        182 eMetAlaIleGlyAsnValLeuGlyTyr------------------------------Al 192

Qy        844 CCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTT 903
              |              |||   |||      :::         |||   |||   :::
Db        192 aThrGlyAlaTyrSerGlyTrpTyrLysValPheProPheSerLeuThrSerSerCysTh 212

Qy        904 CCTCACCTGC------------------------------------GTAGCAGCCAC 924
                :::   |||                                  :::      ||
Db        212 rIleAsnCysAlaAsnLeuLysSerAlaPheTyrIleAspIleIlePheIleIleIleTh 232

Qy        925 ACTGCTGGTGGCTGAGGAGGCAGCGCTG-------------------------GGCCC 957
              |      ::::::      ||||||                           |||||
Db        232 rThrTyrIleSerIleSerAlaAlaLysGluArgProArgIleSerSerGlnAspGlyPr 252

Qy        958 CACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCG 1017
              |            :::|||   :::                  |||   |||
Db        252 oGlnPheSerGluAspGlyThrAlaGln---------SerGlyHisIle---------- 265

Qy       1018 GGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTG 1077
                                          |||   |||         |||
Db        266 ---------------------GluGluAlaPheLeuTrpGluLeuPheGlyThrPheAr 278

Qy       1078 CCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACT 1137
                 :::|||   ::::::      :::   :::               :::|||:::
Db        278 gLeuLeuProGlySerValTrpValIleLeuLeuValThrCysLeuAsnTrpIleGlyTr 298

Qy       1138 CATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCC 1197
                     |||   ||||||   ||||||:::::|||         :::|||   |||   ||
Db        298 pPheProPheIleLeuPheAspThrAspTrpMetGlyArgGluIleTyrGlyGlyGluPr 318

Qy       1198 CAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAG 1257
              |                       :::   |||   :::|||||||||||||||::
Db        318 o-----------------AsnGlnGlyGlnSerTyrSerAspGlyValArgMetGlyAl 332

Qy       1258 CCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCT 1317
              :  ||||||   :::         :::         |||:::::|||::::::||
Db        332 aPheGlyLeuMetMetAsnSerValValLeuGlyIleThrSerValLeuMetGluLysLe 352

Qy       1318 GGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGC 1377
              |  :::   :::|||:::      ::::::
Db        352 uCysArgIleTrpGlySerGlyPheMetTrp-------------------------- 362

Qy       1378 TGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCAC 1437
                                                        |||      |||
Db        363 ----------------------------------------GlyLeuSerAsnIleLeuMe 369

Qy       1438 CGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCG 1497
                     |||   |||:::   ::::::   :::         ::::::   :::
Db        369 tThrIleCysPhePheAlaMetLeuLeuIleThrPheIleAlaLysAsnMetAspTyr-- 388

Qy       1498 GGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGA 1557

Db        388 ------------------------------------------------------------ 388

Qy       1558 CAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACA 1617

Db        388 ------------------------------------------------------------ 388

Qy       1618 CGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGC 1677
                                     |||         |||||||||            |||
Db        389 ------------------GlyThrAsnProProProAsn---------GlyIleValIl 399
```

```
Qy       1678 CTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTG---GTTCC 1734
                   :::   |||      ::::::|||  |||                 |||||
Db        399 eSerAlaLeuIleValPheAlaIleLeuGlyIleProLeuAlaIleThrTyrSerValPr 419

Qy       1735 G--------------------------------GGCCGGGGCATCTGCCTGGACCTCGC 1761
              |                                 |||   |||     |||   :::
Db        419 oTyrAlaLeuValSerThrArgIleGluSerLeuGlyLeuGlyGlnGlyLeuSerMetGl 439

Qy       1762 CATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCAT 1821
                 :::|||:::   |||   ::::::   ||||||      ||||||     |||
Db        439 yValLeuAsnLeuAlaIleValValProGlnValIleValSerLeuGlySerGlyProTr 459

Qy       1822 TGTCCAGCTC---------------AGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGC 1866
                 ||||||                 |||   :::               :::||||||
Db        459 pAspGlnLeuPheGlyGlyGlyAsnSerProAlaPheValValAlaAlaLeuSerAlaPh 479

Qy       1867 AGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTT 1926
                    ||||||:::|||:::              :::         |||||||||   ||
Db        479 eAlaAlaGlyLeuIleAlaLeuIleAlaIleArgArgProArgValAspLysSerArgLe 499

Qy       1927 G 1927
              |
Db        499 u 499
```

RESULT 42

```
CGHU3B
collagen alpha 3(IV) chain precursor, long splice form - human
N;Alternate names: Goodpasture antigen; procollagen alpha 3(IV) chain long splice form
C;Species: Homo sapiens (man)
C;Date: 28-Oct-1994 #sequence_revision 03-Oct-1995 #text_change 22-Jun-1999
C;Accession: A54763; A43928; A44043; A45971; A39786
R;Mariyama, M.; Leinonen, A.; Mochizuki, T.; Tryggvason, K.; Reeders, S.T.
J. Biol. Chem. 269, 23013-23017, 1994
A;Title: Complete primary structure of the human alpha3(IV) collagen chain. Coexpressior
A;Reference number: A54763; MUID:94364994; PMID:8083201
A;Accession: A54763
A;Molecule type: mRNA
A;Residues: 1-1670 <MAR>
A;Cross-references: GB:X80031; NID:g577563; PID:g577564
A;Experimental source: kidney
R;Turner, N.; Mason, P.J.; Brown, R.; Fox, M.; Povey, S.; Rees, A.; Pusey, C.D.
J. Clin. Invest. 89, 592-601, 1992
A;Title: Molecular cloning of the human Goodpasture antigen demonstrates it to be the al
A;Reference number: A43928; MUID:92147878; PMID:1737849
A;Accession: A43928
A;Molecule type: mRNA
A;Residues: 1331-1524,'I',1526-1670 <TUR>
A;Cross-references: GB:M81379
A;Experimental source: kidney
R;Quinones, S.; Bernal, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J.
J. Biol. Chem. 267, 19780-19784, 1992
A;Title: Exon/intron structure of the human alpha 3(IV) gene encompassing the Goodpastur
ction.
A;Reference number: A44043; MUID:93015826; PMID:1400291
A;Accession: A44043
A;Molecule type: DNA; mRNA
A;Residues: 1386-1670 <QUI>
A;Cross-references: GB:M92993; NID:g177895; PIDN:AAA21610.1; PID:g177896
A;Note: sequence extracted from NCBI backbone (NCBIP:115597)
R;Quinones, S.; Bernal, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J.
J. Biol. Chem. 269, 17358, 1994
A;Reference number: A44738; MUID:94274734; PMID:8006044
A;Contents: annotation; erratum; correction to intronic sequence in A44043
R;Bernal, D.; Quinones, S.; Saus, J.
J. Biol. Chem. 268, 12090-12094, 1993
A;Title: The human mRNA encoding the Goodpasture antigen is alternatively spliced.
A;Reference number: A45971; MUID:93280184; PMID:8505332
A;Accession: A45971
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1427-1444 <BER>
A;Note: sequence extracted from NCBI backbone (NCBIP:133363); sequence incorrectly ident
R;Morrison, K.E.; Mariyama, M.; Yang-Feng, T.L.; Reeders, S.T.
Am. J. Hum. Genet. 49, 545-554, 1991
A;Title: Sequence and localization of a partial cDNA encoding the human alpha3 chain
A;Reference number: A39786; MUID:91353570; PMID:1882840
A;Accession: A39786
A;Molecule type: mRNA
A;Residues: 1453-1593,'A',1595-1670 <MOR>
A;Cross-references: GB:S55790; NID:g234418; PIDN:AAB19637.1; PID:g234419
C;Comment: Prolines and lysines at the third position of the tripeptide repeating uni
ed and subsequently O-glycosylated.
C;Comment: In Goodpasture's syndrome, an autoimmune response develops against an epit
C;Genetics:
A;Gene: GDB:COL4A3
A;Cross-references: GDB:128351; OMIM:120070
A;Map position: 2q36-2q37
A;Introns: 1385/1; 1418/1; 1488/1; 1547/2; 1585/3; 1643/2 #status incomplete
A;Note: the alpha 3(IV) and alpha 4(IV) chain genes are encoded on opposite strands w
C;Complex: This minor type IV collagen is thought to form a heterotrimer of two alpha
mong trimer amino-terminal domains (with disulfide and desmosine cross-links), dimeri
er associations in the interrupted helical domain (with disulfide and desmosine cross
C;Function:
A;Description: minor structural component of extracellular basement membrane in kidne
C;Superfamily: collagen alpha 1(IV) chain
C;Keywords: alternative splicing; basement membrane; cell binding; coiled coil; extra
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-1670/Product: collagen alpha 3(IV) chain, long splice form #status predicted <MA
F;29-42/Domain: amino-terminal nonhelical, NH1 <NH1>
F;43-1438/Region: interrupted helical
F;791-793/Region: cell attachment (R-G-D) motif
F;996-998/Region: cell attachment (R-G-D) motif
F;1154-1156/Region: cell attachment (R-G-D) motif
F;1306-1308/Region: cell attachment (R-G-D) motif
F;1345-1347/Region: cell attachment (R-G-D) motif
F;1432-1434/Region: cell attachment (R-G-D) motif
F;1439-1670/Domain: carboxyl-terminal nonhelical, NC1 <NC1>
F;1451-1551/Domain: collagen IV carboxyl-terminal repeat <CT1>
F;1561-1665/Domain: collagen IV carboxyl-terminal repeat <CT2>
F;31,33,39,41,125,422,476,479,682,722,809,1387/Disulfide bonds: interchain #status pr
F;253/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;1460-1548,1493-1551/Disulfide bonds: (or 1460-1551, 1493-1548) #status predicted
F;1505-1511,1616-1622/Disulfide bonds: #status predicted
F;1570-1662,1604-1665/Disulfide bonds: (or 1570-1665, 1604-1662) #status predicted

Alignment Scores:
Pred. No.:              4.08e-11      Length:         1670
Score:                  309.00        Matches:        278
Percent Similarity:     29.78%        Conservative:   74
Best Local Similarity:  23.52%        Mismatches:     397
Query Match:            4.96%         Indels:         435
DB:                     1             Gaps:           59

US-09-759-143-110 (1-3410) x CGHU3B (1-1670)

Qy       3071 TCTGGTCCTGCAGTAGCTCCAAACAGGGTTGTGGAGCTGGTGGGGAAAGTTGGGGGTAGG 3012
                 :::||||||         |||                |||   |||     |||   :::
Db         72 ThrGlyProGln---GlyProLysGlyPheProGlyLeuProGlyLeuThrGlySerLys 90

Qy       3011 GGAAAGTTGGGGGTAGGGGAAATTTGGGCAGTGCCTTCATCAGCCCAGTCCTAGAGAGA 2952
                 |||      |||:::        :::   |||    :::       ||||||   |||
Db         91 GlyValArgGlyIleSerGlyLeuProGlyPheSerGly---SerProGlyLeuProGly 109

Qy       2951 GTAGAGGG---------------------------------------AGTGGAAGTGGG 2931
                    |||                                           |||||||||
Db        110 ThrProGlyAsnThrGlyProTyrGlyLeuValGlyValProGlyCysSerGlySerLys 129

Qy       2930 GGGAACCAGGCTGGGCCAAGAGAAGAGGGGTGGTTAGGGAAGCCGTTGAGACCTGAAGCC 2871
                 |||   |||        |||               |||   ||||||   |||
Db        130 GlyGluGlnGlyPheProGlyLeuProGlyThrLeuGlyTyrPro--------------- 144

Qy       2870 CCACCCTCTACCTTCCTTCAACACCCTAACCTTGGGTAACAGCATTTGGAATTATCATTT 2811

Db        144 ------------------------------------------------------------ 144
```

```
Qy  2810 GGGATGAGTAGAATTTCCAAGGTCCTGGGTTAGGCATTTTGGGGGGCCAGACCCCAGGAG 2751
Db   145 GlyIleProGlyAlaAlaGlyLeuLysGlyGlnLysGlyAlaProAlaLysGluGluAsp 164

Qy  2750 AAGAAGATTCTGGCAATGATCAGCCCA---ATGACCAGCTATCTCAGGGGACCTGATTGT 2694
Db   165 IleGluLeuAspAlaLysGlyAspProGlyLeu-ProGlyAlaProGlyProGlnGlyLe 184

Qy  2693 TGGGGATCCCCCACCCTACCCAAAT---ATTAGACACCAACACAGAAAAGCTAGCAATGG 2637
Db   184 uProGlyProProGlyPheProGlyProValGlyProProGlyProProGlyPhePheGl 204

Qy  2636 ATTCCCTTCTACTTTGTTAAATAAATAAGTTAAATATTTAAATGCCTGTGTCTCTGTGAT 2577
Db   204 yPhePro------------------------------------------GlyAlaMe 209

Qy  2576 GGCAACAGAAGGACCAACAGGCCAC---------------ATCCTGATAAAAGGTAAGAG 2532
Db   209 tGlyProArgGlyProLysGlyHisMetGlyGluArgValIleGlyHisLysGlyGluAr 229

Qy  2531 GGGGTGGATCAGCAAAAAGACAGTGCTGTGGGCTGAGGGGACCTGGTTCTTGTGTGTTG 2472
Db   229 gGly-------------------------------------------------------- 230

Qy  2471 CCCCTCAGGACTCTTCCCCTACAAATAAGTCATATGTTCAAATCCCATGGAGGAGTGTTT 2412
Db   230 ------------------------------------------------------------ 230

Qy  2411 CATCCTAGAAACTCCCATGCAAGAGCTACATTAAACGAAGCTGCAGGTTAAGGGGCTTAG 2352
Db   231 ---------------------------------------------ValLysGlyLeuTh 235

Qy  2351 A------GATGGGAAACCAGGTGACTGAGTT------TATTCAGCTCCCAAAAACCCTTC 2304
Db   235 rGlyProProGlyProProGlyThrValIleValThrLeuThrGlyProAspAsnArgTh 255

Qy  2303 TCTAGGTGTGTCTCAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGC 2244
Db   255 rAspLeuLysGlyGluLysGlyAsp---------LysGlyAlaMetGlyGluProGlyPr 272

Qy  2243 AGAGTCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGA 2184
Db   272 oPro--------------------GlyProSerGlyLeuPro----------------- 279

Qy  2183 AACCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGAC 2124
Db   280 ------------------------------GlyGluSerTyrGlySerGluLysGlyAl 289

Qy  2123 GCCCCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGAGCC 2064
Db   289 aProGlyAspProGlyLeuGln------------GlyLysProGlyLysAspGlyValPr 305

Qy  2063 ACATTACTTTGGCAGC---------------AACAGAAACTGGCGGCCAGCCCGGCAGCC 2019
Db   305 oGlyPheProGlySerGluGlyValLysGlyAsnArgGlyPheProGlyLeuMetGlyGl 325

Qy  2018 CCATGGG---GCTAACAGGAGCGGGGAGCTGGGACCCAGTGAGGCAGGCCCTCCACCCCA 1962
Db   325 uAspGlyIleLysGlyGlnLysGlyAspIleGlyProProGlyPheArgGlyPro----- 343

Qy  1961 ATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCT 1902
Db   344 ---------------------------------------------ThrGluTyrTyrAs 348

Qy  1901 GTGTAGCAAAGTAAATGGCGACCAGAC-------------CCAGGCCTGCGGCAGACACC 1855
Db   348 pThrTyrGlnGluLysGlyAspGluGlyThrProGlyProProGlyProArg-------- 365

Qy  1854 ATATAGGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCATAAACAGGGATGGGGCC 1795
Db   366 ------------------GlyAlaArgGlyProGlnGlyProSerGlyProProGlyVa 379

Qy  1794 ACCTGGGACAGCAGGAAGGCACTATCCAGGATGGCGAGGTCCAGGCAGATGCCCCGGCCC 1735
Db   379 lProGlySerProGlySerSerArgProGlyLeuArgGly------------------- 392

Qy  1734 GGAACCACCCTGGCCTCG----GTGGGCTCACCC---ACCACCACACGTACGGAGACATC 1682
Db   393 -AlaProGlyTrpProGlyLeuIleGlySerLysGlyGluArgGlyArgProGlyLysAs 412

Qy  1681 ACAGGCAGAGGCCCCGCAGAGCGCGGGTGGAGGTGGGAGCAGGCCACTGCCT------CC 1628
Db   412 pAlaMetGlyThrProGlySerProGlyCysAlaGlySerProGlyLeuProGlySerPr 432

Qy  1627 AGCACCCACGTGTCCATTAGGGAAG----------------------------------- 1603
Db   432 oGlyProProGlyProProGlyAspIleGlyPheArgLysGlyProProGlyAspHisGl 452

Qy  1602 ----------------GGAGCTCCAGGCTTA-------------GGGCCTGGCAGGAAGCT 1571
Db   452 yLeuProGlyTyrLeuGlySerProGlyIleProGlyValAspGlyPro-LysGlyGluP 472

Qy  1570 GGTCATCAGGCTGTCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAA 1511
Db   472 roGlyLeuLeuCysThrGlnCysProTyrIleProGlyProProGlyLeu---ProGlyL 491

Qy  1510 CACCTGCTTCTCCCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGA 1451
Db   491 euProGlyLeuHisGlyValLysGlyIlePro---GlyArgGlnGlyAlaAlaGly---- 508

Qy  1450 GAAGGTGAACCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCA 1391
Db   509 ----------------------LeuLysGlySerProGlySerProGlyAsnThrGly- 520

Qy  1390 TGTGGCACCGGCAGCCACAGGGAAAGCTGCCACACTGGCCA------------------- 1350
Db   521 -------------------------LeuProGlyPheProGlyPheProGlyAlaGlnG 532

Qy  1349 ------------------------------AATAGACTGCTCGAGTGCCGAATCGCTG 1322
Db   532 lyAspProGlyLeuLysGlyGluLysGlyGluThrLeuGlnProGluGlyGlnValGlyV 552

Qy  1321 CACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACAGCCC 1262
Db   552 alProGlyAspProGlyLeuArgGlyGlnProGlyArgLysGlyLeuAspGly---IleP 571

Qy  1261 CAGGCTGCCCATCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGC 1202
Db   571 roGlyThrProGlyValLys---------GlyLeuProGlyProLysGlyGluLeuAlaL 588

Qy  1201 TCTGGGCACGCCCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGT 1142
Db   588 euSerGlyGluLysGlyAspGlnGlyProProGlyAspPro------------------ 601

Qy  1141 CATGAGTGCCATCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCAT 1082
Db   602 -------------------------------------------------GlySerP 604

Qy  1081 GCGGCAGCACAGCTGGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGCG 1022
Db   604 roGlySer---ProGlyProAlaGlyProAlaGlyProProGlyTyrGlyProGlnGlyG 623

Qy  1021 GGCCCGGCATGGACAGCAGT-------------------------------GGGGCGA 995
Db   623 luProGlyLeuGlnGlyThrGlnGlyValProGlyAlaProGlyProProGlyGluAlaG 643

Qy   994 CAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCTCCTCAGC 935
Db   643 lyProArgGlyGluLeuSerValSerThrProValProGlyProProGlyProProGlyP 663

Qy   934 CACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGCA 875
Db   663 roPro-----------------------GlyHisProGlyProGlnGlyProProGlyI 675

Qy   874 CTCCT-----------------------------------------------------CCTG 866
```

```
Db        675 leProGlySerLeuGlyLysCysGlyAspProGlyLeuProGlyProAspGlyGluProG 695

Qy        865 GGTGCCCAGGTAGGGGG-------------------------------------CCAG 845
              ||  ||||||  |||                                      ||||
Db        695 lyIleProGlyIleGlyPheProGlyProProGlyProLysGlyAspGlnGlyPheProG 715

Qy        844 GGCAC-----------TGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGCC 797
              ||              ||||||  :::            |||  ||||||   |
Db        715 lyThrLysGlySerLeuGlyCysProGlyLysMetGlyGluProGlyLeuProGlyLysP 735

Qy        796 CCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGTCCG------- 744
              ||                               ||||||    |||
Db        735 ro--------------------------GlyLeuProGlyAlaLysGlyGluProAlaV 746

Qy        743 -------------GGTCCCGGAAGA-----------GGTCAGAGAGCA---------- 720
                           |||||||||              |||:::|||
Db        746 alAlaMetProGlyGlyProGlyThrProGlyPheProGlyGluArgGlyAsnSerGlyG 766

Qy        719 ----------------------GGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAA 683
                                    |||  |||          |||||||||:::
Db        766 luHisGlyGluIleGlyLeuProGlyLeuProGlyLeuProGlyThrProGlyAsnGluG 786

Qy        682 GTCCAGCAGCCCCACGC------CCAGGATGAGCAGTGCCAGCTCCAGGGGCCTGGGATC 629
                      ||||||     ||||||       |||  ||||||        |
Db        786 lyLeuAspGlyProArgGlyAspProGlyGlnProGlyPro---ProGlyGluGlnGlyP 805

Qy        628 CGGGCACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGAT 569
              ||                 |||   :::             |||||||:::   |||
Db        805 ro-----------------ProGlyArgCysIleGluGlyProArgGlyAlaGlnGlyL 819

Qy        568 GCCCAAGGACAGTGCCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTC 509
                |||                         |||                      |||
Db        819 euPro-----------------------GlyLeuAsnGlyLeuLysGlyGlnGlnGlyA 831

Qy        508 ACTGGCTGAGCCTAGGAGCGGGACACAGACCAGGC---CCAGCACTGGACCAATGCCCAG 452
                                 |||       ||||||  |||   |||||||:::
Db        831 rgArgGlyLysThrGlyProLysGlyAspProGlyIleProGlyLeuAspArgSerGlyP 851

Qy        451 CACCATGGTCATGAACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCACATAGGT 392
                 |||        ||||||     |||          |||
Db        851 heProGlyGluThrGlySerProGlyIleProGlyHisGlnGlyGluMetGly------- 868

Qy        391 GATGCCTGCGGCCAAACACACCTCCAGGCCAAAGG--------------------TTAG 353
                                 |||   |||||||||                   :::
Db        869 ------------------ProLeuGlyGlnArgGlyTyrProGlyAsnProGlyIleL 882

Qy        352 CAGGTTGACCAGCAAGAGCTGGGCTTTCCGGTGCCGCAGCAGGCGGCTCACCCACAGCCT 293
                |||  |||        |||:::  |||            |||  |||  |||
Db        882 euGlyProProGlyGluAspGlyValIleGlyMetMet------GlyPheProGlyAlaI 900

Qy        292 CTGGACCATAGTGGG---CCAGGCGGGTAGGGCTCAGGGGGCCGTTCAGGCACTCCAGAA 236
                ||||||  |||  ||||||   |||:::  |||       |||:::|||
Db        900 leGlyPro-ProGlyProProGlyAsnProGlyThrProGlyHisArgGlySerProGly 919

Qy        235 CTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAACTG 176
               :::                                   |||      |||
Db        920 Ile----------------------------------ProGlyValLysGlyGln--- 926

Qy        175 CCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGCTTCTCAG 116
                 |||   |||||||||       |||          |||
Db        927 ---ArgGlyThrProGlyAlaLysGlyGluGln-----------GlyAspLysGlyAsn 941

Qy        115 CCCATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTCATCACTCAGA 56
              |||        |||:::   :::   |||:::    ||||||                ::
Db        942 Pro------Gly-ProSerGluIleSerHisValIleGlyAsp---------LysGlyGl 956

Qy         55 TCCTGGC---CGAGGCGCGCGGCTGTCACCCGGAGCCAGCGCGTGCAGGCTGGTTCCC 1
              :||||||    :::|||          :::||||||            |||   |||||
Db        956 uProGlyLeuLysGlyPheAlaGlyAsnProGlyGluLysGlyAsnArgGlyValPro 975
```

RESULT 43
CGBO7S
collagen alpha 1(III) chain - bovine

```
C;Species: Bos primigenius taurus (cattle)
C;Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 07-May-1999
C;Accession: A02862; A38001; A38002; A38003; A38004; A38005; S71946
R;Fietzek, P.P.; Allmann, H.; Rauterberg, J.; Henkel, W.; Wachter, E.; Kuehn, K.
Hoppe-Seyler's Z. Physiol. Chem. 360, 809-820, 1979
A;Title: The covalent structure of calf skin type III collagen. I. The amino acid seq
A;Reference number: A02862; MUID:80026026; PMID:488906
A;Accession: A02862
A;Molecule type: protein
A;Residues: 1-242 <FIE>
R;Dewes, H.; Fietzek, P.P.; Kuehn, K.
Hoppe-Seyler's Z. Physiol. Chem. 360, 821-832, 1979
A;Title: The covalent structure of calf skin type III collagen. II. The amino acid se
A;Reference number: A38001; MUID:80026027; PMID:488907
A;Accession: A38001
A;Molecule type: protein
A;Residues: 243-422 <DEW1>
R;Bentz, H.; Fietzek, P.P.; Kuehn, K.
Hoppe-Seyler's Z. Physiol. Chem. 360, 833-840, 1979
A;Title: The covalent structure of calf skin type III collagen. III. The amino acid s
A;Reference number: A38002; MUID:80026028; PMID:488908
A;Accession: A38002
A;Molecule type: protein
A;Residues: 423-571 <BEN>
R;Lang, H.; Glanville, R.W.; Fietzek, P.P.; Kuehn, K.
Hoppe-Seyler's Z. Physiol. Chem. 360, 841-850, 1979
A;Title: The covalent structure of calf skin type III collagen. IV. The amino acid se
A;Reference number: A38003; MUID:80026029; PMID:488909
A;Accession: A38003
A;Molecule type: protein
A;Residues: 572-808 <LAN>
R;Dewes, H.; Fietzek, P.P.; Kuehn, K.
Hoppe-Seyler's Z. Physiol. Chem. 360, 851-860, 1979
A;Title: The covalent structure of calf skin type III collagen. V. The amino acid seq
A;Reference number: A38004; MUID:80026030; PMID:488910
A;Accession: A38004
A;Molecule type: protein
A;Residues: 809-947 <DEW2>
R;Allmann, H.; Fietzek, P.P.; Glanville, R.W.; Kuehn, K.
Hoppe-Seyler's Z. Physiol. Chem. 360, 861-868, 1979
A;Title: The covalent structure of calf skin type III collagen. VI. The amino acid se
A;Reference number: A38005; MUID:80026031; PMID:488911
A;Accession: A38005
A;Molecule type: protein
A;Residues: 948-1049 <ALL>
A;Experimental source: skin
R;Henkel, W.
Biochem. J. 318, 497-503, 1996
A;Title: Cross-link analysis of the C-telopeptide domain from type III collagen.
A;Reference number: S71946; MUID:96404897; PMID:8809038
A;Accession: S71946
A;Molecule type: protein
A;Residues: 87-106;1017-1029;1037-1049 <HEN>
C;Comment: Prolines at the third position of the tripeptide repeating unit (G-X-Y) ar
C;Comment: The type III collagen molecule is a trimer of identical chains, linked to
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo
C;Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxypr
F;1-1049/Product: collagen alpha 1(III) chain #status experimental <CAB>
F;1-14/Region: amino-terminal nonhelical telopeptide
F;15-1040/Region: helical
F;587-589/Region: cell attachment (R-G-D) motif
F;752-754/Region: cell attachment (R-G-D) motif
F;875-877/Region: cell attachment (R-G-D) motif
F;878-880/Region: cell attachment (R-G-D) motif
F;935-937/Region: cell attachment (R-G-D) motif
F;1041-1049/Region: carboxyl-terminal nonhelical telopeptide
F;95,107,119,938,950/Modified site: 5-hydroxylysine (Lys) #status experimental
F;107,950/Modified site: allysine (Lys) #status predicted
F;107/Binding site: carbohydrate (Lys) (covalent) #status experimental
F;1040,1041/Disulfide bonds: interchain #status predicted

Alignment Scores:
Pred. No.:             4.69e-11       Length:       1049
```

```
Score:                     308.00        Matches:        298
Percent Similarity:        31.09%        Conservative:   63
Best Local Similarity:     25.67%        Mismatches:     399
Query Match:               4.80%         Indels:         401
DB:                        1             Gaps:           65

US-09-759-143-110 (1-3410) x CGBO7S (1-1049)

Qy       26 CCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTGATGAGACGTGTCCCCACTGAGG 85
Db      134 ProGlyGluAspGlyAlaProGlyPro-----------------MetGlyProArg-Gl 147

Qy       86 TGCCCCACAGCAGCAGGTGTTG-------------------------------------- 107
Db      147 yAlaProGlyGluArgGlyArgProGlyLeuProGlyAlaAlaGlyAlaArgGlyAsnAs 167

Qy      108 ----AGCATGGGCTGAGAAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGG 163
Db      167 pGlyAlaArgGlySerAspGlyGlnProGlyProProGly-----------ProProGl 183

Qy      164 C------TGATTCCTAGGCAGTTGGCGGCAGCAAGGA--------------GGAGAGGC 202
Db      183 yThrAlaGlyPheProGlySerProGlyAlaLysGlyGluValGlyProAlaGlySerPr 203

Qy      203 CGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAGTGCCTGAACGGCCCCCTGA 262
Db      203 oGlySerSerGlyAlaProGlyGlnArgGlyGluProGlyProGlnGlyHisAlaGlyAl 223

Qy      263 GCCCTACCCGCCTGGCCCACTATGGT-----------CCAGAGGCTGTGGGTGAG---C 307
Db      223 aProGlyProProGlyPro-ProGlySerAspGlySerProGlyGlyLysGlyGluMetG 243

Qy      308 CGCCTGCTGCGGCACCGGAAAGCCCAGCTCTT---------GCTGGTCAACCTGCTAACC 358
Db      243 lyProAlaGlyIleProGlyAlaProGlyLeuIleGlyAlaArgGlyProProGlyProP 263

Qy      359 TTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAA 418
Db      263 roGlyThrAsnGlyValProGlyGlnArg-----------------GlyAlaAlaGlyG 277

Qy      419 GTGGGGGTAGA---GGAGAAGTTCATGACCATGGT---GCTGGGCATTGGTCCAGTGCTG 472
Db      277 luProGlyLysAsnGlyAlaLysGlyAspProGlyProArgGlyGluArgGlyGluAlaG 297

Qy      473 GG---CCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTAT 529
Db      297 lySerProGlyIleAla------------------------------------------ 302

Qy      530 GGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTC 589
Db      303 -----GlyProLysGlyGluAspGlyLysAsp---GlySerProGlyGluProGlyAlaA 320

Qy      590 ATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTG 649
Db      320 snGlyLeuProGlyAlaAlaGlyGluArgGlyValProGlyPheArgGlyProAlaGlyA 340

Qy      650 GCACTGCTCATCCTGGG------------------------------------------ 666
Db      340 laAsnGlyLeuProGlyGluLysGlyProProGlyAspArgGlyGlyProGlyProAlaG 360

Qy      667 -----CGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTG 721
Db      360 lyProArgGlyValAlaGly------GluProGlyArgAsnGlyLeuProGlyGlyProG 378

Qy      722 CTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTC 781
Db      378 lyLeuArgGlyIleProGly---SerProGlyGlyProGly------------------ 390

Qy      782 ATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGT 841
Db      391 --------SerAsnGlyLysProGlyProProGlySerGlnGlyGluThrGlyArg---- 406

Qy      842 GCCCTGGCCCCCTACCTGGGCACCCAGGA--------------GGAGTGCCTCTTTGGC 886
Db      407 --ProGlyPro---ProGlySerProGlyProArgGlyGlnProGlyValMetGlyPheP 425

Qy      887 CTGCT------------------------------------------------------ 891
Db      425 roGlyProLysGlyAsnAspGlyAlaProGlyLysAsnGlyGluArgGlyGlyProGlyG 445

Qy      892 -----CACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCT---GGTGGCTGAGGAG 943
Db      445 lyProGlyProGlnGlyProAlaGlyLysAsnGlyGluThrGlyProGlnGlyProProG 465

Qy      944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
Db      465 lyProThrGlyPro---SerGlyAspLysGlyAspThrGlyProProGlyProGlnGlyL 484

Qy     1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCG---GAACCTGGGCGCCCTGCTTCCCCGG 1060
Db      484 euGlnGlyLeuProGlyThrSerGlyProProGlyGluAsnGlyLysProGlyGluProG 504

Qy     1061 CTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTG 1120
Db      504 ly--------------ProLysGlyGluAlaGlyAlaProGlyIleProGly------- 516

Qy     1121 TGCAGCTGGATGGCACTCATGACCTTCACGCTGTT---TTACACGGATTTCGTGGGCGAG 1177
Db      517 ----------GlyLysGlyAspSerGlyAlaProGlyGluArgGlyProProGlyAlaG 533

Qy     1178 GGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGAT 1237
Db      533 lyGlyProProGly---ProArgGlyGlyAlaGlyProProGlyProGluGly------G 550

Qy     1238 GAAGGCGTTCGGATGGGCAGCCTGGG--------------GCTGTTCCTGCAGTGCGCC 1282
Db      550 lyLysGlyAlaAlaGlyProProGlyProProGlySerAlaGlyThrProGlyLeuGlnG 570

Qy     1283 ATCTCCCTGGTCTTCTC---------TCTGGTCATGGACCGGCTGGTGCAGCGATTCGGC 1333
Db      570 lyMetProGlyGluArgGlyGlyProGlyGlyProGlyProLysGlyAspLysGlyGluP 590

Qy     1334 ACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGC 1393
Db      590 roGlySerSer------GlyValAspGlyAla--------------------------P 599

Qy     1394 CTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCA 1453
Db      599 roGlyLysAspGlyProArgGlyProThr----------------------------G 609

Qy     1454 GCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTC 1513
Db      609 lyProIleGlyProProGlyProAlaGlyGlnProGlyAspLysGlyGluSerGlyAlaP 629

Qy     1514 CTGCC------------CAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGC 1561
Db      629 roGlyValProGlyIleAlaGlyProArgGly--------------GlyProGlyGluA 644

Qy     1562 CTGATGACCAGCTTCCTGCCAGGCCCTAAG------CCTGGAGCTCCCTTCCCTAATGGA 1615
Db      644 rgGlyGluGlnGlyPro-ProGlyProAlaGlyPheProGlyAlaProGlyGlnAsnGly 663

Qy     1616 CACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCA-----------------CCTCCACCC 1657
Db      664 GluProGlyAlaLysGlyGluArgGlyAlaProGlyGluLysGlyGluGlyGlyProPro 683

Qy     1658 GCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACC 1717
Db      684 Gly-AlaAlaGlyProAlaGlyGlySerGlyPro---------------AlaGlyProPr 698

Qy     1718 G-----------------------------AGGCCAGGGTGGTTCCGGGCCGGGGCATC 1747
Db      698 oGlyProGlnGlyValLysGlyGluArgGlySerPro-GlyGlyProGlyAlaAlaGlyP 718

Qy     1748 TGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCA---------GGTGGCC 1798
```

```
Db        718 heProGlyGlyArgGlyPro-----------ProGlyProProGlySerAsnGlyAsnP 734

Qy       1799 CCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTG 1858
              ||  |||                      |||   :::                   |||
Db        734 roGlyProProGlySerSerGlyAlaProGlyLysAsp-----------------GlyP 748

Qy       1859 TCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAG 1918
                          |||||||||                ::::      :::       |||
Db        748 roProGlyProProGlySerAsn-GlyAlaProGlySerProGlyIleSerGlyProLys 767

Qy       1919 AGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGC 1978
                |||                          |||        |||  |||
Db        768 GlyAsp--------------------------SerGlyProProGlyGluArgGlyAla 778

Qy       1979 CTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCA 2038
                                 |||  |||     |||  |||  ||||||      |||
Db        779 -----------------ProGlyProGlnGlyProProGlyAlaProGly------Pro 790

Qy       2039 GTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCG 2098
              :::   :::|||                 |||:::   |||
Db        791 LeuGlyIleAlaGly-----------LeuThrGlyAlaArg------------------ 800

Qy       2099 TAGCTGCACAGCTGGGGCTGGGGCGTCCCTCTCCTCTCCCCAGTCTCTAGGGCTGCC 2158
                             ||||||   |||     ::: |||   :::   |||  ||
Db        801 --------------GlyLeuAlaGlyProProGlyMet-ProGlyAlaArgGlySerPr 815

Qy       2159 TGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGG------CCAGA 2212
              |       :::     ||||||   :::                |||:::        |||
Db        815 oGlyProGlnGlyIleLysGlyGluAsn--------------GlyLysProGlyProSe 830

Qy       2213 AGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAA 2272
               |||   :::                                |||||||||   |||
Db        830 rGlyGlnAsnGlyGluArgGlyProProGlyProGlnGlyLeuProGlyLeuAlaGlyTh 850

Qy       2273 CAGCTAGCCTCCTAGTTGAGAC-----------------------ACACCTAGAGAAGG 2308
               :::      |||     |||                               |||:::||
Db        850 rAlaGlyGluProGlyArgAspGlyAsnProGlySerAspGlyLeuProGlyArgAspGl 870

Qy       2309 GTTTTTGGGAGCTGAATAAACTCAG------------TCACCTGGTTTCCCATCT----- 2351
              |        ||||||:::       :::            |||||||||   |||
Db        870 yAlaProGlyAlaLysGlyAspArgGlyGluAsnGlySerProGlyAlaProGlyAlaPr 890

Qy       2352 ------------------------CTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAG 2386
                                      :::   |||
Db        890 oGlyHisProGlyProProGlyProValGlyProAla---------------------- 902

Qy       2387 CTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTA 2446

Db        902 ------------------------------------------------------------ 902

Qy       2447 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 2506
                       |||||||||   :::           |||           |||
Db        903 -------GlyLysSerGlyAspArgGlyGluThrGlyProAlaGlyPro---------- 916

Qy       2507 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 2566
                                |||         |||         |||   |||||
Db        917 ----------------SerGlyAlaProGlyProAlaGlySerArgGlyProProGlyPr 931

Qy       2567 T-----------------------------------------TCTGTTGCCATCACAGA 2584
              |                                         ::::::  |||
Db        931 oGlnGlyProArgGlyAspLysGlyGluThrGlyGluArgGlyAlaMetGlyIleLysGl 951

Qy       2585 GACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCT 2644
                    :::   |||                              |||||||||
Db        951 yHisArgGlyPhe-----------------------------ProGlyAsnPro----- 959

Qy       2645 AGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTC 2704
                                                   |||:::|||        ||
Db        960 ----------------------------------------GlyAlaPro------GlySe 964

Qy       2705 CCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGG 2764
              ||||        |||||||||                          :::||||||
Db        964 rProGlyProAlaGlyHisGlnGly--------------------AlaValGlySerPr 977

Qy       2765 CCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCA 2824
               |||                        :::|||  |||     |||
Db        977 oGlyProAlaGlyProArgGlyProValGlyProSerGlyProPro-------------- 992

Qy       2825 AATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCA 2884
                                                 |||:::     |||||||||||
Db        993 --------------------------------GlyLys-AspGlyAlaSerGly---- 999

Qy       2885 ACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCACTTCCACTCC 2944
                               ||||||       :::|||||| ||||||
Db       1000 --------------HisProGlyProIleGlyPro-ProGlyProArgGlyAsnArgGly 1014

Qy       2945 CCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCA 3004
                                                     |||      |||     |||
Db       1015 GluArgGlySerGluGly-----------------SerProGlyHisProGlyGlnPro 1028

Qy       3005 ACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGT 3049
                 |||   |||         |||    ||||||   ||||||
Db       1029 GlyProProGlyPro---------ProGlyAlaProGlyProCys 1040
```

RESULT 44
B40333
collagen alpha 1(II) chain precursor - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: B40333
R;Su, M.W.; Suzuki, H.R.; Bieker, J.J.; Solursh, M.; Ramirez, F.
J. Cell Biol. 115, 565-575, 1991
A;Title: Expression of two nonallelic type II procollagen genes during Xenopus laevis
A;Reference number: A40333; MUID:92011898; PMID:1918153
A;Accession: B40333
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1486 <SUA>
A;Cross-references: GB:M63595
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo
C;Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix
F;37-96/Domain: von Willebrand factor type C repeat homology <VWC>
F;1258-1486/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

```
Alignment Scores:
Pred. No.:               4.69e-11     Length:          1486
Score:                   308.00       Matches:         277
Percent Similarity:      33.50%       Conservative:    54
Best Local Similarity:   28.04%       Mismatches:      362
Query Match:             4.95%        Indels:          296
DB:                      1            Gaps:            61

US-09-759-143-110 (1-3410) x B40333 (1-1486)

Qy       2569 GAAGGACCAACAGGCCACATCCTGATAAAAGGTAAGAGGGGGGTGGATCAGCAAAAAGAC 2510
               :::|||||||:::|||                   :::|||:::||||||   |||:::   :::
Db        130 GlnGlyProSerGlyGluGlnGlySerArgGlyGluArgGly---AspLysGlyGluLys 148

Qy       2509 AGTGCTGTGGGCTGAGGGGACCTGGTTCTTGTGTGTTGCCCCTCAGGACTCTTCCCCTAC 2450
                 |||   |||                          |||   |||
Db        149 GlyAlaProGly--------------------------ProArgGly----------- 155

Qy       2449 AAATAAGTCATATGTTCAAATCCCATGGAGGAGTGTTTCATCCTAGAAACTCCCATGCAA 2390

Db        155 ------------------------------------------------------------ 155

Qy       2389 GAGCTACATTAAACGAAGCTGCAGGTTAAGGGGCTTAGAGATGGGAAACCAGGTGACTGA 2330
                                                   |||||||||:::||||||
Db        156 ------------------------------------ArgAspGlyGluProGly------ 161

Qy       2329 GTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCTCAACTAGGAGGCTAGCTGTTA 2270
                       |||   ||||||
Db        162 ---------ThrProGlyAsnProGly--------------------------------- 167

Qy       2269 ACCCTGAGCCTGGGTAATCCACCTGCAGAGTCCCCGCATTCCAGTGCATGGAGCCCTTCT 2210
                :::          ||||||                 :::   :::|||
```

```
Db      168 ProValGlyProProGlyProProGlyLeuGlyGlyAsnPheAlaAlaGlnMetThrGly 187

Qy     2209 GGCCTCCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGCCTCCAGTCAGGC------ 2156
            |||                        :::         |||  |||  |||
Db      188 GlyPhe------------------AspGluLysAlaGlyGlyAlaGlnMetGlyValMet 201

Qy     2155 ---AGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCCCCAGCCC---------CCAGCT 2108
                |||                |||||| |||  |||              |||
Db      202 GlnGlyProMetGlyProMetGlyProArgGlyProProGlyProThrGlyAlaProGly 221

Qy     2107 GTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGC 2048
               |||                ||||||        ||||||        |||:::
Db      222 ProGln------------GlyPheGlnGlyAsnProGlyGluProGlyGluProGlyAla 237

Qy     2047 AACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGGA 1988
                        |||       |||      :::   :::   |||:::   |||
Db      238 GlyGlyProMetGlyProArgGlyProProGlyProSerGlyLysProGlyAspAspGly 257

Qy     1987 CCCAGT--------------------GAGGCAGGCCCTCCA---CCCCAATGTGCTGGA 1952
             :::                      |||   |||||||||   ||||||  |||
Db      258 GluAlaGlyLysProGlyLysSerGlyGluArgGlyProProGlyProGlnGlyAlaArg 277

Qy     1951 AGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTT------GTCAAA----------- 1910
             |||             |||       |||       ||||||
Db      278 GlyPhePro------------GlyThrProGlyLeuProGlyValLysGlyHisArgGly 293

Qy     1909 TACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCCTGCGGCAGACACCATATA 1850
             |||                   |||   :::      |||  |||
Db      294 TyrProGlyLeuAspGlyAlaLysGly---GluAlaGlyAlaAlaGlyAlaLysGlyGlu 312

Qy     1849 GGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCATAAACAGGGATGGGGCCA---- 1794
             |||       ||||||   ||||||              |||      |||||||||
Db      313 GlyGly--AlaThrGlyGluAlaGlySerProGlyPro---------MetGlyProArgG 329

Qy     1793 -----CCTGGGACAGCAGGAAGG------------CACTATCCAGGATGGCGAGGTCCAG 1751
                  ||||||                              |||   ||||||
Db      329 lyLeuProGlyGluArgGlyArgProGlySerSerGlyAlaAlaGlyAlaArgGlyAsnA 349

Qy     1750 GCAGATGCCCCGGCCCGGAA---CCACCCTGGCCTCG-------GTGGGCTCACCCACCA 1701
                    |||||||||      ||||||   |||             |||:::|||
Db      349 spGlyLeuProGlyProAlaGlyProProGlyProValGlyProAlaGlyAlaProGlyP 369

Qy     1700 CCACACGTACGGAGACATCACAGGCAGAGGCC---------------------CCGCAGA 1662
                             |||:::   ||||||                    |||:::
Db      369 heProGlyAlaProGlySerLysGlyGluAlaGlyProThrGlyAlaArgGlyProGluG 389

Qy     1661 GCGCGGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGGG---A 1605
              |||   |||                              |||      |||
Db      389 lyAlaGlnGlyProArgGlyGluSerGlyThrProGlySerProGlyProAlaGlyAlaS 409

Qy     1604 AGGGAGCTCCAGGCTTAGGG------CCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCAC 1551
              |||   ||||||              ||||||   |||   |||||||||   |||
Db      409 erGlyAsnProGlyThrAspGlyIleProGlyAlaLys-GlySerSerGlyGlyPro--- 427

Qy     1550 TGCTAGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGT 1491
                                ||||||   ||||||   |||       ||||||
Db      428 -----------------------GlyIle---AlaGlyAlaProGlyPheProGlyPro 438

Qy     1490 AGAGGGAGGCCAGTGTGTAGGGCA---------------GGATCTGCAGGGCTGAGAAGG 1446
             ||||||   |||         :::                |||     :::|||   ||
Db      439 ArgGlyProProGlyProGlnGlyAlaThrGlyProLeuGlyPro-LysGlyGluThrGl 458

Qy     1445 TGAACCCGGTGAGGGCGGCTGAAGCTGTCACCACGGC------CACACTGTGGGACAGGC 1392
             |:::||||||         |||                   :::|||   ||
Db      458 yAspProGlyValAlaGlyPheLysGlyGluGlnGlyProLysGlyGluIleGlySerAl 478

Qy     1391 ATGTGGCACCGGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGA---G 1335
             |          ||| |||  |||  |||           |||     ||||||
Db      478 aGlyProGlnGly-AlaProGlyProAlaGlyGluGluGlyLysArgGlyAlaArgGlyG 498

Qy     1334 TGCCGAATCGCTGCA---CCAGCCGGTCCATGACCAGAGAGAAGA-------------- 1293
              |||   ||||||   |||   ||||||      ||||||
Db      498 luProGly-AlaAlaGlyProAsnGlyProProGlyGluArgGlyAlaProGlyAsnArg 517

Qy     1292 ------CCAGGGAGATGGCGCACTGCAGGAACAGCCCCAGGCTGCCCATCCGAACGCCTT 1239
                  ||||||:::              ||||||      |||  |||  ||||||
Db      518 GlyPheProGlyGlnAspGlyLeuAlaGly---ProLysGlyAlaProGlyGluArgGly 536

Qy     1238 CATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTACAGCC 1179
               :::   ||||||:::                                  |||
Db      537 ValProGlyLeuGlyGlyProLys------------------GlyGlyAsnGlyAspPro 550

Qy     1178 CCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGCTGCACA 1119
               ||||||   |||                |||                      |||
Db      551 GlyArgProGlyGluProGlyLeuProGlyAlaArg--------------GlyLeuThr 565

Qy     1118 GCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACAGCTGGTGCAGCC 1059
              :::|||      ||||||             |||   |||      :::||||||:::
Db      566 GlyArgProGlyAspAlaGlyProGln---GlyLysValGly---ProSerGlyAlaSer 583

Qy     1058 GGGGAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGG-----------CCCGGCATG 1011
            ||||||   ||||||||||||         |||   |||            ||||||:::
Db      584 GlyGluAspGlyArgProGlyProProGlyProGlnGlyAlaArgGlyGlnProGlyVal 603

Qy     1010 ------------GACAGCAGTGGGGCGACA------------AGGAGGGGGCCGACAGCC 975
                           ||||||                    :::   |||
Db      604 MetGlyPheProGlyProLysGlyAlaAsnGlyGluProGlyLysAlaGlyGluLysGly 623

Qy      974 CTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCT------------------------- 942
            |||:::   |||                ||||||
Db      624 LeuValGlyAlaProGlyLeuArgGlyLeuProGlyLysAspGlyGluThrGlySerGln 643

Qy      941 ---CCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGC 885
               |||    ||||||                                   |||  ||||||
Db      644 GlyProAsnGlyProAla---------------------------GlyProAlaGly 653

Qy      884 CAAAGAGGCACT------CCTCCTGGGTGCCCAGGTAGGGGGCCAGGGCACTGGTGTCCC 831
            :::||||||            |||||||||                        |||
Db      654 GluArgGlyGluGlnGlyProProGly---------------------------Pro 663

Qy      830 AGTCAATGGCAGGCA---GGAGGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCAT 774
            |||    :::|||          |||    ||||||||||||     :::
Db      664 SerGlyPheGlnGlyLeuProGlyProProGlySerProGlyGlu-------------- 678

Qy      773 AGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGT---CCCGGAAGAGGTCAGAGAGCA 720
                       :::|||||||||   |||   ||||||      |||
Db      679 ---GlyGlyLysProGlyAspGlnGlyValProGlyGluAlaGlyAlaProGlyLeuVal 697

Qy      719 GGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGC------ 666
            ||||||   |||          ||||||  |||            :::   |||:::
Db      698 GlyProArgGlyGluArgGlyPheProGlyGluArgGlySerSerGlyProGlnGlyLeu 717

Qy      665 -----------------CCAGGATGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGC 624
                              ||||||                      ||||||      |||
Db      718 GlnGlyProArgGlyLeuProGly------------ThrProGlyThrAspGlyProLys 733

Qy      623 ACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCA 564
                            ||||||   |||   |||           |||:::  |||  |||
Db      734 Gly-----------AlaSerGlyProSerGly---ProAsnGlyAlaGlnGlyProPro 748

Qy      563 ------AGGACAGTGCCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCC------ 516
                  :::   :::|||         |||       |||   |||
Db      749 GlyLeuGlnGlyMetProGlyGluArgGlyAlaAlaGlyIleSerGlyProLysGlyAsp 768

Qy      515 ---------AGTGGTCACTGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTG 465
                     :::|||         ||||||          :::   |||
Db      769 ArgGlyAspThrGlyGluLysGlyProGluGlyAlaSerGlyLysAspGlySerArg--- 787

Qy      464 GACCAATGCCCAGCACCATGGTCATGAACTTCTCCTCTA---CCCCCACTTCCAGCAGCA 408
                                     |||  |||:::   ||||||   ||||||
Db      788 -------------------GlyLeuThrGlyProIleGlyProProGlyProAlaGly 800

Qy      407 GAGGCGGCACATAGGTGATGCCTGCGGCCAAACACACCTCCAGGC-------------- 363
               :::                  |||:::   |||||||||
Db      801 ProAsnGlyGluLysGlyGluSerGlyProSerGlyProProGlyIleValGlyAlaArg 820
```

```
Qy        362 ----------------------------------------CAAAGGTTAGCAGGTTGA 345
                                                              ||||||
Db        821 GlyAlaProGlyAspArgGlyGluAsnGlyProProGlyProAlaGlyPheAlaGlyPro 840

Qy        344 CCAGCAAGAGCTGGGCTTTCCGGTGCCGCAGCA--------------------------- 312
               |||          |||   ||||||
Db        841 ProGlyAlaAspGlyGlnSerGlyLeuLysGlyAspGlnGlyGluSerGlyGlnLysGly 860

Qy        311 ---GGCGGCTCACCCACAGCCTCTGGACCATAGTGGG---CCAGGCGGGTAGGGCTCAGG 258
                  |||:::|||            |||||| ||||||   ||||||     |||
Db        861 AspAlaGlyAlaProGlyProGlnGlyPro-SerGlyAlaProGlyProGlnGlyProTh 880

Qy        257 GGGCCGTTCAGGCACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTC 198
               |||      |||   |||:::        |||            |||      |||
Db        880 rGlyValPheGly---ProLysGlyAlaArgGlyAlaGlnGlyProAlaGlyAlaThrGl 899

Qy        197 TCCTCCTTGCTGCCGCCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTT 138
                 |||    ||||||        :::|||                 :::      |||
Db        899 yPhePro-GlyAlaAlaGlyArgValGlyThrProGlyProAsnGlyAsnProGlyProP 919

Qy        137 TGGTGCCGGTCCAGCTTCTCAGCCCATGCTCAACA-----------CCTGCTGCTGTGG 90
                 |||                  |||   |||             |||     |||
Db        919 roGlyPro---------------ProGlySerAlaGlyLysGluGlyProLysGlyValA 934

Qy         89 GGCACCTCAGTGGGGACACGTCTCATCACTCAGATCCTGGC---CGAGGCGCGCGGCTGT 33
                    :::|||                    |||||||||   :::||||||     :
Db        934 rgGlyAspAlaGlyProProGlyArgAlaGlyAspProGlyLeuGlnGlyAlaAlaGlyA 954

Qy         32 CACCCGGA 25
              ::||||||
Db        954 laProGly 956
```

RESULT 45
F96741
probable sucrose transport protein F17M19.4 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C;Accession: F96741
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: F96741
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-512 <STO>
A;Cross-references: GB:AE005173; NID:g6978914; PIDN:AAF34306.1; GSPDB:GN00141
C;Genetics:
A;Gene: F17M19.4
A;Map position: 1
C;Superfamily: common tobacco sucrose transport protein

```
Alignment Scores:
Pred. No.:               5.39e-11      Length:         512
Score:                   307.00        Matches:        128
Percent Similarity:      42.15%        Conservative:   92
Best Local Similarity:   24.52%        Mismatches:     208
Query Match:             4.78%         Indels:         95
DB:                      2             Gaps:           19

US-09-759-143-110 (1-3410) x F96741 (1-512)

Qy        289 CCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAA 348
               ||||||   :::|||:::|||          :::|||        :::::::::   |||
Db         22 ProGluAspLeuGlyGlnPro---------SerPro-LeuArgLysIleIleSerValAl 38

Qy        349 CCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCT 408
                 :::      |||:::::      |||      ::::::    :::   |||
Db         38 aSerIleAlaAlaGlyValGlnPheGlyTrpAlaLeuGlnLeuSerLeuLeuThrProTy 58

Qy        409 GCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGT 468
                 :::     :::|||:::      |||:::   :::    :::         ||||||::
Db         58 rIleGlnLeuLeuGlyIleProHisLysTrpSerSerTyrMetTrpLeuCysGlyProIl 78

Qy        469 GCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTA 528
               :   |||::::::         |||::::::|||     ||||||           |||::
Db         78 eSerGlyMetIleValGlnProIleValGlyTyrHisSerAspArgCysGluSerArgPh 98

Qy        529 TGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCT 588
               :||||||||||||||||||||       |||                 :::   :::|||:::|||||
Db         98 eGlyArgArgArgProPheIleAlaAlaGlyValAlaLeuValAlaValSerValPheLe 118

Qy        589 CATCCCAAGGGCCGGCTGGCTAGCAGGGCTG-------------------CTGTGCCC 627
               ||||         |||:::   |||   :::                       |||
Db        118 uIle---------GlyPheAlaAlaAspMetGlyHisSerPheGlyAspLysLeuGluAs 135

Qy        628 GGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTG 687
                  |||          :::   :::   :::   |||         ||||||
Db        135 nLysValArgThrArgAlaIleIleIlePheLeuThrGlyPheTrpPheLeuAspValAl 155

Qy        688 TGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGG---GACCC 744
                       |||      |||   |||:::||||||              |||
Db        155 aAsnAsnThrLeuGlnGlyProCysArgAlaPheLeuAlaAspLeuAlaAlaGlyAspAl 175

Qy        745 GGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCT 804
                    |||   |||   :::   ::::::|||    ::::::::::|||      ||
Db        175 aLysLysThrArgValAlaAsnAlaCysPheSerPhePheMetAlaValGlyAsnValLe 195

Qy        805 GGGCTAC----------------------------CTCCTGCCTGCCATTGACTGGGA 834
               |||||||                              :::  |||
Db        195 uGlyTyrAlaAlaGlySerTyrThrAsnLeuHisLysMetPheProPheThr------Me 213

Qy        835 CACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCAC 894
               |||   |||        |||      :::   |||   |||         :::||
Db        213 tThrLysAlaCysAspIleTyrCysAlaAsnLeuLysThrCysPhePheLeuSerIleTh 233

Qy        895 CCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGG 954
               ||||:::   |||            ::::::||| |||:::::::::    ::::::|||:
Db        233 rLeuLeu---LeuIleValThrPheSerSerLeu-TrpTyrValLysAspLysGlnTrpS 252

Qy        955 CCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATG 1014
               ::||||||      |||
Db        252 erProProGlnGlyAspLys----------------------------GluGluLysT 262

Qy       1015 CCGGGCCCGCTTGGCTTTCCGGAAC----CTGGGCGCCCTGCTTCCCCGGCTGCACCAGC 1070
                 :::      |||          ||||||:::
Db        262 hrSerSerLeuPhePhePheGlyGluIlePheGlyAlaValArg--------------- 276

Qy       1071 TGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGA 1130
                        |||   |||   :::         |||   :::      :::   :::|||:
Db        277 --------HisMetLysArgProMetValMetLeuLeuIleValThrValIleAsnTrpI 294

Qy       1131 TGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGG 1190
               ::|||      |||   |||:::   ||||||:::::|||         :::|||   |
Db        294 leAlaTrpPheProPheIleLeuTyrAspThrAspTrpMetGlyArgGluValTyrGlyG 314

Qy       1191 GCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGA 1250
               ||        ::::::                  :::::::::   ||||||:::||||||:::
Db        314 ly------AsnSerAspGlyAspGluArgSerLysLysLeuTyrAspGlnGlyValGlnA 332

Qy       1251 TGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGG 1310
                   |||:::||||||||||              :::         ||||||   ::::
Db        332 laGlyAlaLeuGlyLeuMetPheAsnSerIleLeuLeuGlyPheValSerLeuGlyValG 352

Qy       1311 ACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCC 1370
              ::   :::   ::::::   |||      |||      |||                |||
Db        352 luSerIleGlyArgLysMetGly---GlyAlaLysArgLeuTrpGlyCysValAsnPheI 371

Qy       1371 CTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT------- 1423
```

```
                 :::|||  |||          |||:::  :::|||  |||      :::
Db         371 leLeuAlaIleGly---------LeuAlaMetThrValLeuValThrLysSerAlaGluH 388

Qy        1424 ----------------------------------TCAGCCGCCCTCACCGGGTTCACCT 1448
                                                 |||:::   :::           |
Db         388 isHisArgGluIleAlaGlyProLeuAlaGlyProSerSerGlyIleLysAlaGlyValP 408

Qy        1449 TCTCAGCCCTGCAGATCCTG-----------------------------CCCTACACAC 1478
               |||||          :::|||                            |||:::   |
Db         408 heSerLeuPheThrValLeuGlyIleProLeuAlaIleThrTyrSerIleProPheAlaL 428

Qy        1479 TGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTG 1538
               ||||||||::::::                                     |||
Db         428 euAlaSerIlePheSerThrAsnSerGlyAlaGlyGlnGlyLeuSerLeuGlyValLeuA 448

Qy        1539 GAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCT---- 1594
                    |||                ::::::   ||||||     ||||||
Db         448 snIleAlaIleCysIleProGlnMetIleValSerPheSerSerGlyProLeuAspAlaG 468

Qy        1595 --------GGAGCTCCCTTCCCTAAT---------GGACACGTGGGTGCTGGAGGCAGTG 1637
                       |||        |||:::          |||   :::   |||      ||||
Db         468 lnPheGlyGlyGlyAsnLeuProSerPheValValGlyAlaIleAlaAlaAlaValSerG 488

Qy        1638 GCCTGCTC--------------------CCACCTCCA------CCCGCGCTCTGCGGGG 1670
               ||:::|||                    |||||||||      ||||||:::   ||||
Db         488 lyValLeuAlaLeuThrValLeuProSerProProProAspAlaProAlaMetSerGlyA 508

Qy        1671 CC 1672
               ||
Db         508 la 508
```

Search completed: June 9, 2003, 22:35:30
Job time : 309.5 secs

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

```
Run on:          June  9, 2003, 22:32:24 ; Search time 159 Seconds
                                            (without alignments)
                                            4428.293 Million cell updates/sec

Title:           US-09-759-143-110
Perfect score:   6418
Sequence:        1 gggaaccagcctgcacgcgc..........aaaaaaataaaaaaaaaaaa 3410

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters:      784170

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09759143/runat_05062003_111321_23172/app_query.fasta_1.3591
-DB=Published_Applications_AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=45 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09759143_@CGN_1_1_92_@runat_05062003_111321_23172
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       Published_Applications_AA:*
                 1:  /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
                 2:  /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
                 3:  /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
                 4:  /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
                 5:  /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
                 6:  /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
                 7:  /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
                 8:  /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
                 9:  /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
                 10:  /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
                 11:  /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
                 12:  /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
                 13:  /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
                 14:  /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2861	44.6	553	9	US-09-232-880-113	Sequence 113, App
	2	2861	44.6	553	9	US-10-012-896-113	Sequence 113, App
	3	2861	44.6	553	9	US-09-895-793-113	Sequence 113, App
	4	2861	44.6	553	9	US-09-895-814-113	Sequence 113, App
	5	2861	44.6	553	9	US-10-010-940-113	Sequence 113, App
	6	2861	44.6	553	10	US-09-745-288-101	Sequence 101, App
	7	2861	44.6	553	10	US-09-838-785-2	Sequence 2, Appli
	8	2861	44.6	553	10	US-09-759-143-113	Sequence 113, App
	9	2861	44.6	553	10	US-09-780-669-113	Sequence 113, App
	10	2861	44.6	553	10	US-09-030-606-113	Sequence 113, App
	11	2861	44.6	553	10	US-09-822-827-113	Sequence 113, App
	12	2861	44.6	553	10	US-09-115-453-113	Sequence 113, App
	13	2861	44.6	1079	9	US-09-895-793-947	Sequence 947, App
	14	2861	44.6	1079	10	US-09-822-827-947	Sequence 947, App
	15	1696	26.4	359	9	US-09-895-793-974	Sequence 974, App
	16	1696	26.4	359	10	US-09-822-827-974	Sequence 974, App
	17	1420.5	22.1	371	9	US-10-012-896-708	Sequence 708, App
	18	1420.5	22.1	371	9	US-09-895-793-708	Sequence 708, App
	19	1420.5	22.1	371	9	US-09-895-814-708	Sequence 708, App
	20	1420.5	22.1	371	10	US-09-759-143-708	Sequence 708, App
	21	1420.5	22.1	371	10	US-09-780-669-708	Sequence 708, App
	22	1420.5	22.1	371	10	US-09-822-827-708	Sequence 708, App
	23	1403.5	21.9	400	9	US-10-012-896-852	Sequence 852, App
	24	1403.5	21.9	400	9	US-09-895-793-852	Sequence 852, App
	25	1403.5	21.9	400	9	US-09-895-814-852	Sequence 852, App
	26	1403.5	21.9	400	10	US-09-759-143-852	Sequence 852, App
	27	1403.5	21.9	400	10	US-09-780-669-852	Sequence 852, App
	28	1403.5	21.9	400	10	US-09-822-827-852	Sequence 852, App
	29	947	14.8	355	9	US-10-012-896-1011	Sequence 1011, Ap
	30	452	7.0	84	9	US-10-010-940-571	Sequence 571, App
	31	426	6.6	123	9	US-10-012-896-706	Sequence 706, App
	32	426	6.6	123	9	US-09-895-793-706	Sequence 706, App
	33	426	6.6	123	9	US-09-895-814-706	Sequence 706, App
	34	426	6.6	123	10	US-09-759-143-706	Sequence 706, App
	35	426	6.6	123	10	US-09-780-669-706	Sequence 706, App
	36	426	6.6	123	10	US-09-822-827-706	Sequence 706, App
c	37	356.5	5.7	837	9	US-10-096-986-103	Sequence 103, App
c	38	347	5.6	1464	9	US-10-216-705-21	Sequence 21, Appl
c	39	343	5.5	1464	9	US-10-060-036-159	Sequence 159, App
c	40	343	5.5	1464	9	US-10-171-311-36	Sequence 36, Appl
	41	323	5.0	837	9	US-10-096-986-103	Sequence 103, App
c	42	316	5.1	1806	10	US-09-919-497-56	Sequence 56, Appl
	43	312	4.9	58	9	US-10-012-896-547	Sequence 547, App
	44	312	4.9	58	9	US-09-895-793-547	Sequence 547, App
	45	312	4.9	58	9	US-09-895-814-547	Sequence 547, App

ALIGNMENTS

```
RESULT 1
US-09-232-880-113
; Sequence 113, Application US/09232880
; Publication No. US20020182596A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT: Dillon, Davin C.
;  APPLICANT: Mitcham, Jennifer Lynn
;  TITLE OF INVENTION: COMPOUNDS FOR IMMUNODIAGNOSIS OF
;  TITLE OF INVENTION: PROSTATE CANCER AND METHODS FOR THEIR USE
;  FILE REFERENCE: 210121.428C6
;  CURRENT APPLICATION NUMBER: US/09/232,880
;  CURRENT FILING DATE: 1999-01-15
;  NUMBER OF SEQ ID NOS: 338
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 113
;   LENGTH: 553
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-232-880-113

Alignment Scores:
Pred. No.:               3.32e-151     Length:         553
Score:                   2861.00       Matches:        553
Percent Similarity:      100.00%       Conservative:   0
Best Local Similarity:   100.00%       Mismatches:     0
Query Match:             44.58%        Indels:         0
DB:                      9             Gaps:           0
```

```
US-09-759-143-110 (1-3410) x US-09-232-880-113 (1-553)

Qy      284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy      344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy      404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy      464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy      524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy      584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy      644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy      704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy      764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy      824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy      884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy      944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy     1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy     1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy     1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy     1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy     1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy     1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy     1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy     1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy     1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy     1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy     1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy     1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy     1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy     1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy     1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy     1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
            |||||||||||||||||||||||||||||||||||||||
Db      541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 2
US-10-012-896-113
; Sequence 113, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Wantanabe, Yoshihiro
;  APPLICANT:  Meagher, Madeleine Joy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
```

```
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012,896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 113
;   LENGTH: 553
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-012-896-113

Alignment Scores:
Pred. No.:              3.32e-151     Length:        553
Score:                  2861.00       Matches:       553
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            44.58%        Indels:        0
DB:                     9             Gaps:          0

US-09-759-143-110 (1-3410) x US-10-012-896-113 (1-553)

Qy    284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy    344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy    404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy    464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy    524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy    584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy    644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy    704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy    764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy    824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy    884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy    944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy   1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy   1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy   1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy   1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy   1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy   1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy   1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy   1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy   1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy   1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy   1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy   1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy   1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy   1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy   1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy   1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
          |||||||||||||||||||||||||||||||||||||||
Db    541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 3
US-09-895-793-113
; Sequence 113, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT: Dillon, Davin C.
;  APPLICANT: Mitcham, Jennifer L.
;  APPLICANT: Harlocker, Susan L.
;  APPLICANT: Jiang, Yuqiu
;  APPLICANT: Kalos, Michael D.
;  APPLICANT: Retter, Marc W.
;  APPLICANT: Stolk, John A.
```

```
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.534C2
;  CURRENT APPLICATION NUMBER: US/09/895,793
;  CURRENT FILING DATE:  2001-06-29
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 113
;   LENGTH: 553
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-895-793-113

Alignment Scores:
Pred. No.:               3.32e-151      Length:         553
Score:                   2861.00        Matches:        553
Percent Similarity:      100.00%        Conservative:   0
Best Local Similarity:   100.00%        Mismatches:     0
Query Match:             44.58%         Indels:         0
DB:                      9              Gaps:           0

US-09-759-143-110 (1-3410) x US-09-895-793-113 (1-553)

Qy       284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy       344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy       404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy       464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy       524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy       584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy       644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy       704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy       764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy       824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy       884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy       944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy      1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy      1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy      1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy      1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy      1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy      1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy      1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy      1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy      1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy      1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy      1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy      1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy      1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy      1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy      1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy      1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
```

```
              ||||||||||||||||||||||||||||||||||||||
Db        541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 4
US-09-895-814-113
; Sequence 113, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C26
;  CURRENT APPLICATION NUMBER: US/09/895,814
;  CURRENT FILING DATE:  2001-06-29
;  NUMBER OF SEQ ID NOS: 990
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 113
;   LENGTH: 553
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-895-814-113

Alignment Scores:
Pred. No.:              3.32e-151     Length:        553
Score:                  2861.00       Matches:       553
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            44.58%        Indels:        0
DB:                     9             Gaps:          0

US-09-759-143-110 (1-3410) x US-09-895-814-113 (1-553)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy         1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy         1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy         1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy         1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
                |||||||||||||||||||||||||||||||||||||||
Db          541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 5
US-10-010-940-113
; Sequence 113, Application US/10010940
; Publication No. US20030088062A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan Louise
;  APPLICANT:  Jiang Yuqui
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Kalos, Michael
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Retter, Mark
;  APPLICANT:  Solk, John
;  APPLICANT:  Day, Craig
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427D3
;  CURRENT APPLICATION NUMBER: US/10/010,940
;  CURRENT FILING DATE:  2001-12-05
;  NUMBER OF SEQ ID NOS: 575
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 113
;   LENGTH: 553
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-10-010-940-113

Alignment Scores:
Pred. No.:              3.32e-151      Length:         553
Score:                  2861.00        Matches:        553
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            44.58%         Indels:         0
DB:                     9              Gaps:           0

US-09-759-143-110 (1-3410) x US-10-010-940-113 (1-553)

Qy          284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy          344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy          404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy          464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy          524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy          584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy          644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy          704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy          764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy          824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy          884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy          944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy         1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy         1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy         1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy         1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy         1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy         1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy         1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy         1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy         1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy         1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440
```

```
Qy      1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy      1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy      1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy      1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy      1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy      1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
             |||||||||||||||||||||||||||||||||||||||
Db       541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 6
US-09-745-288-101
; Sequence 101, Application US/09745288
; Patent No. US20010018058A1
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  TITLE OF INVENTION:  COMPOUNDS FOR IMMUNOTHERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
;  FILE REFERENCE:  210121.446D1
;  CURRENT APPLICATION NUMBER: US/09/745,288
;  CURRENT FILING DATE:  2000-12-19
;  NUMBER OF SEQ ID NOS:  101
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 101
;   LENGTH: 553
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-745-288-101

Alignment Scores:
Pred. No.:              3.32e-151     Length:        553
Score:                  2861.00       Matches:       553
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            44.58%        Indels:        0
DB:                     10            Gaps:          0

US-09-759-143-110 (1-3410) x US-09-745-288-101 (1-553)

Qy       284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy       344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy       404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy       464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy       524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy       584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy       644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy       704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy       764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy       824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy       884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy       944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy      1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy      1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy      1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy      1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy      1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy      1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy      1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy      1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy      1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy      1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy      1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460
```

```
Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||
Db        541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553
```

RESULT 7
US-09-838-785-2
; Sequence 2, Application US/09838785
; Patent No. US20020009455A1
; GENERAL INFORMATION:
; APPLICANT: Lau, Ted
; APPLICANT: Lin, Rick
; APPLICANT: Parkes, Debbie
; APPLICANT: Parry, Gordon
; APPLICANT: Schneider, Douglas
; APPLICANT: Steinbrecher, Renate
; APPLICANT: Van Heuit, Pam T
; APPLICANT: Wu, John
; TITLE OF INVENTION: DNA Encoding a No. US20020009455A1el PROST 03
; FILE REFERENCE: 51831AUSM1
; CURRENT APPLICATION NUMBER: US/09/838,785
; CURRENT FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/200,065
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-838-785-2

```
Alignment Scores:
Pred. No.:              3.32e-151      Length:        553
Score:                  2861.00        Matches:       553
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            44.58%         Indels:        0
DB:                     10             Gaps:          0
```

US-09-759-143-110 (1-3410) x US-09-838-785-2 (1-553)

```
Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440
```

```
Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||
Db        541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553
```

```
RESULT 8
US-09-759-143-113
; Sequence 113, Application US/09759143
; Patent No. US20020022248A1
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C23
;  CURRENT APPLICATION NUMBER: US/09/759,143
;  CURRENT FILING DATE:  2001-01-12
;  NUMBER OF SEQ ID NOS:  934
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 113
;   LENGTH: 553
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-759-143-113

Alignment Scores:
Pred. No.:              3.32e-151      Length:         553
Score:                  2861.00        Matches:        553
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            44.58%         Indels:         0
DB:                     10             Gaps:           0

US-09-759-143-110 (1-3410) x US-09-759-143-113 (1-553)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy      1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy      1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy      1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy      1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy      1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy      1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy      1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy      1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
             |||||||||||||||||||||||||||||||||||||||
Db       541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553
```

RESULT 9
US-09-780-669-113
; Sequence 113, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780,669
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 113
;   LENGTH: 553
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-780-669-113

```
Alignment Scores:
Pred. No.:              3.32e-151      Length:          553
Score:                  2861.00        Matches:         553
Percent Similarity:     100.00%        Conservative:    0
Best Local Similarity:  100.00%        Mismatches:      0
Query Match:            44.58%         Indels:          0
DB:                     10             Gaps:            0

US-09-759-143-110 (1-3410) x US-09-780-669-113 (1-553)

Qy       284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy       344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy       404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy       464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy       524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy       584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy       644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy       704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy       764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy       824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy       884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy       944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy      1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy      1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy      1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy      1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320
```

```
Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||
Db        541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 10
US-09-030-606-113
; Sequence 113, Application US/09030606
; Patent No. US20020081580A1
;  GENERAL INFORMATION:
;    APPLICANT:  Xu, Jiangchun
;    APPLICANT:  Dillon, Davin C.
;    TITLE OF INVENTION:  COMPOUNDS FOR IMMUNODIAGNOSIS OF PROSTATE CANCER AND METHODS F
;    NUMBER OF SEQUENCES:  224
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/030,606
;      FILING DATE:  25-FEB-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.428C3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  113:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  553 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
US-09-030-606-113

Alignment Scores:
Pred. No.:              3.32e-151     Length:         553
Score:                  2861.00       Matches:        553
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            44.58%        Indels:         0
DB:                     10            Gaps:           0

US-09-759-143-110 (1-3410) x US-09-030-606-113 (1-553)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220
```

```
Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||
Db        541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553
```

RESULT 11
US-09-822-827-113
; Sequence 113, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-822-827-113

```
Alignment Scores:
Pred. No.:               3.32e-151     Length:         553
Score:                   2861.00       Matches:        553
Percent Similarity:      100.00%       Conservative:   0
Best Local Similarity:   100.00%       Mismatches:     0
Query Match:             44.58%        Indels:         0
DB:                      10            Gaps:           0

US-09-759-143-110 (1-3410) x US-09-822-827-113 (1-553)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||
Db        541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 12
US-09-115-453-113
; Sequence 113, Application US/09115453B
; Patent No. US20020090372A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT: Dillon, Davin C.
;  TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND
;  TITLE OF INVENTION: METHODS FOR THEIR USE
;  FILE REFERENCE: 210121.427C4
;  CURRENT APPLICATION NUMBER: US/09/115,453B
;  CURRENT FILING DATE: 1998-07-14
;  NUMBER OF SEQ ID NOS: 228
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 113
;   LENGTH: 553
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-115-453-113

Alignment Scores:
Pred. No.:                3.32e-151     Length:          553
Score:                    2861.00       Matches:         553
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              44.58%        Indels:          0
DB:                       10            Gaps:            0

US-09-759-143-110 (1-3410) x US-09-115-453-113 (1-553)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
```

```
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||
Db        541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553
```

RESULT 13
US-09-895-793-947
; Sequence 947, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 947
; LENGTH: 1079
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-793-947

```
Alignment Scores:
Pred. No.:               3.26e-151     Length:         1079
Score:                   2861.00       Matches:        553
Percent Similarity:      100.00%       Conservative:   0
Best Local Similarity:   100.00%       Mismatches:     0
Query Match:             44.58%        Indels:         0
DB:                      9             Gaps:           0

US-09-759-143-110 (1-3410) x US-09-895-793-947 (1-1079)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 546

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 566

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 586

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 606

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 626

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 646

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 666

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 686

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 706

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       707 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 726

Qy       884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       727 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 746

Qy       944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       747 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 766

Qy      1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       767 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 786

Qy      1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       787 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 806

Qy      1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       807 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 826

Qy      1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       827 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 846

Qy      1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       847 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 866

Qy      1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       867 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 886

Qy      1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       887 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 906

Qy      1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       907 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 926

Qy      1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       927 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 946

Qy      1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       947 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 966

Qy      1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       967 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 986

Qy      1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       987 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 1006

Qy      1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1007 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 1026

Qy      1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1027 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 1046

Qy      1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1047 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 1066

Qy      1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
             |||||||||||||||||||||||||||||||||||||||
Db      1067 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 1079
```

RESULT 14
US-09-822-827-947
; Sequence 947, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 947
;   LENGTH: 1079
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-822-827-947

```
Alignment Scores:
Pred. No.:              3.26e-151      Length:          1079
Score:                  2861.00        Matches:         553
Percent Similarity:     100.00%        Conservative:    0
Best Local Similarity:  100.00%        Mismatches:      0
Query Match:            44.58%         Indels:          0
DB:                     10             Gaps:            0

US-09-759-143-110 (1-3410) x US-09-822-827-947 (1-1079)

Qy       284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       527 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 546

Qy       344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       547 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 566

Qy       404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       567 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 586

Qy       464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       587 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 606

Qy       524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       607 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 626

Qy       584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       627 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 646

Qy       644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       647 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 666

Qy       704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       667 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 686

Qy       764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       687 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 706

Qy       824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       707 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 726

Qy       884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        727 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 746

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 766

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 786

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 806

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 826

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 846

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        847 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 866

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 886

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 906

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 926

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 946

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        947 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 966

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 986

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 1006

Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 1026

Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 1046

Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 1066

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||
Db       1067 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 1079
```

RESULT 15
US-09-895-793-974
; Sequence 974, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 974
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-793-974

```
Alignment Scores:
Pred. No.:               1.35e-86    Length:          359
Score:                   1696.00     Matches:         320
Percent Similarity:      100.00%     Conservative:    0
Best Local Similarity:   100.00%     Mismatches:      0
Query Match:             26.43%      Indels:          0
DB:                      9           Gaps:            0
```

US-09-759-143-110 (1-3410) x US-09-895-793-974 (1-359)

```
Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

RESULT 16
US-09-822-827-974
; Sequence 974, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.534C1
;  CURRENT APPLICATION NUMBER: US/09/822,827
;  CURRENT FILING DATE:  2001-03-28
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 974
;   LENGTH: 359
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-822-827-974

Alignment Scores:
Pred. No.:               1.35e-86      Length:        359
Score:                   1696.00       Matches:       320
Percent Similarity:      100.00%       Conservative:  0
Best Local Similarity:   100.00%       Mismatches:    0
Query Match:             26.43%        Indels:        0
DB:                      10            Gaps:          0

US-09-759-143-110 (1-3410) x US-09-822-827-974 (1-359)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

RESULT 17
US-10-012-896-708
; Sequence 708, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
```

```
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Wantanabe, Yoshihiro
;  APPLICANT:  Meagher, Madeleine Joy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C27
;  CURRENT APPLICATION NUMBER: US/10/012,896
;  CURRENT FILING DATE:  2001-12-10
;  NUMBER OF SEQ ID NOS: 1011
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 708
;   LENGTH: 371
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-012-896-708

Alignment Scores:
Pred. No.:              2.57e-71      Length:        371
Score:                  1420.50       Matches:       275
Percent Similarity:     88.58%        Conservative:  12
Best Local Similarity:  84.88%        Mismatches:    23
Query Match:            22.13%        Indels:        14
DB:                     9             Gaps:          1

US-09-759-143-110 (1-3410) x US-10-012-896-708 (1-371)

Qy        347 AACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCT 406
              :::|||:::         |||   :::   |||:::      :::|||:::
Db          9 SerLeuValProLeuProLeuAlaLeuTyrLeuSerGlnProLeuThrHisThrThrSer 28

Qy        407 CTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCA 466
              ||||||                                          ||||||||||||
Db         29 LeuLeu------------------------------------AlaGlyIleGlyPro 35

Qy        467 GTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGC 526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         36 ValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArg 55

Qy        527 TATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTT 586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         56 TyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPhe 75

Qy        587 CTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAG 646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 LeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGlu 95

Qy        647 CTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACT 706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 LeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThr 115

Qy        707 CCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTAC 766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        116 ProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyr 135

Qy        767 TCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATT 826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 SerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIle 155

Qy        827 GACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGC 886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 AspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPheGly 175

Qy        887 CTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCA 946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 LeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAla 195

Qy        947 GCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGC 1006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 AlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCys 215

Qy       1007 TGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCAC 1066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 CysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHis 235

Qy       1067 CAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGC 1126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 GlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSer 255

Qy       1127 TGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTAC 1186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 TrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeuTyr 275

Qy       1187 CAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTT 1246
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 GlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGlyLys 295

Qy       1247 CGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTC 1306
                 ::: ||||||            |||   :::      |||   :::::|||    ::
Db        296 AlaLeu-AlaAlaSerArgGlyTrpCysGlySerArgProProGluThrThrLeuGlyAl 315

Qy       1307 ATGGACCGGC 1316
              :   :::|||
Db        315 aValSerGly 318

RESULT 18
US-09-895-793-708
; Sequence 708, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.534C2
;  CURRENT APPLICATION NUMBER: US/09/895,793
;  CURRENT FILING DATE:  2001-06-29
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 708
;   LENGTH: 371
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-895-793-708

Alignment Scores:
```

```
Pred. No.:                 2.57e-71       Length:        371
Score:                     1420.50        Matches:       275
Percent Similarity:        88.58%         Conservative:  12
Best Local Similarity:     84.88%         Mismatches:    23
Query Match:               22.13%         Indels:        14
DB:                        9              Gaps:          1

US-09-759-143-110 (1-3410) x US-09-895-793-708 (1-371)

Qy        347 AACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCT 406
              :::|||:::           |||   :::   |||:::      :::|||:::
Db          9 SerLeuValProLeuProLeuAlaLeuTyrLeuSerGlnProLeuThrHisThrThrSer 28

Qy        407 CTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCA 466
              ||||||                                          ||||||||||||
Db         29 LeuLeu------------------------------------------AlaGlyIleGlyPro 35

Qy        467 GTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGC 526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         36 ValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArg 55

Qy        527 TATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTT 586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         56 TyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPhe 75

Qy        587 CTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAG 646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 LeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGlu 95

Qy        647 CTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACT 706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 LeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThr 115

Qy        707 CCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTAC 766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        116 ProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyr 135

Qy        767 TCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATT 826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 SerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIle 155

Qy        827 GACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGC 886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 AspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPheGly 175

Qy        887 CTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCA 946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 LeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAla 195

Qy        947 GCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGC 1006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 AlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCys 215

Qy       1007 TGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCAC 1066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 CysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHis 235

Qy       1067 CAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGC 1126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 GlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSer 255

Qy       1127 TGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTAC 1186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 TrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeuTyr 275

Qy       1187 CAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTT 1246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 GlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGlyLys 295

Qy       1247 CGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTC 1306
                 :::  ||||||          |||   :::       |||   :::::|||   ::
Db        296 AlaLeu-AlaAlaSerArgGlyTrpCysGlySerArgProProGluThrThrLeuGlyAl 315

Qy       1307 ATGGACCGGC 1316
               :   :::|||
Db        315 aValSerGly 318
```

RESULT 19
US-09-895-814-708
; Sequence 708, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 708
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-814-708

```
Alignment Scores:
Pred. No.:                 2.57e-71       Length:        371
Score:                     1420.50        Matches:       275
Percent Similarity:        88.58%         Conservative:  12
Best Local Similarity:     84.88%         Mismatches:    23
Query Match:               22.13%         Indels:        14
DB:                        9              Gaps:          1

US-09-759-143-110 (1-3410) x US-09-895-814-708 (1-371)

Qy        347 AACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCT 406
              :::|||:::           |||   :::   |||:::      :::|||:::
Db          9 SerLeuValProLeuProLeuAlaLeuTyrLeuSerGlnProLeuThrHisThrThrSer 28

Qy        407 CTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCA 466
              ||||||                                          ||||||||||||
Db         29 LeuLeu------------------------------------------AlaGlyIleGlyPro 35

Qy        467 GTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGC 526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         36 ValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArg 55

Qy        527 TATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTT 586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         56 TyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPhe 75

Qy        587 CTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAG 646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 LeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGlu 95
```

```
Qy      647 CTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACT 706
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       96 LeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThr 115

Qy      707 CCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTAC 766
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      116 ProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyr 135

Qy      767 TCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATT 826
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      136 SerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIle 155

Qy      827 GACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGC 886
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      156 AspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPheGly 175

Qy      887 CTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCA 946
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      176 LeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAla 195

Qy      947 GCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGC 1006
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      196 AlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCys 215

Qy     1007 TGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCAC 1066
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      216 CysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHis 235

Qy     1067 CAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGC 1126
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      236 GlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSer 255

Qy     1127 TGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTAC 1186
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      256 TrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeuTyr 275

Qy     1187 CAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTT 1246
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      276 GlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGlyLys 295

Qy     1247 CGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTC 1306
               :::  |||||         |||   :::     |||   :::::|||   ::
Db      296 AlaLeu-AlaAlaSerArgGlyTrpCysGlySerArgProProGluThrThrLeuGlyAl 315

Qy     1307 ATGGACCGGC 1316
            :   :::|||
Db      315 aValSerGly 318

RESULT 20
US-09-759-143-708
; Sequence 708, Application US/09759143
; Patent No. US20020022248A1
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C23
;  CURRENT APPLICATION NUMBER: US/09/759,143
;  CURRENT FILING DATE:  2001-01-12
;  NUMBER OF SEQ ID NOS:  934
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 708
;   LENGTH: 371
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-759-143-708

Alignment Scores:
Pred. No.:                2.57e-71      Length:         371
Score:                    1420.50       Matches:        275
Percent Similarity:       88.58%        Conservative:   12
Best Local Similarity:    84.88%        Mismatches:     23
Query Match:              22.13%        Indels:         14
DB:                       10            Gaps:           1

US-09-759-143-110 (1-3410) x US-09-759-143-708 (1-371)

Qy      347 AACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCT 406
             :::|||:::          |||   :::   |||:::      :::|||:::
Db        9 SerLeuValProLeuProLeuAlaLeuTyrLeuSerGlnProLeuThrHisThrThrSer 28

Qy      407 CTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCA 466
            |||||                                         ||||||||||||
Db       29 LeuLeu---------------------------------------AlaGlyIleGlyPro 35

Qy      467 GTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGC 526
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       36 ValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArg 55

Qy      527 TATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTT 586
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       56 TyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPhe 75

Qy      587 CTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAG 646
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       76 LeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGlu 95

Qy      647 CTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACT 706
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       96 LeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThr 115

Qy      707 CCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTAC 766
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      116 ProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyr 135

Qy      767 TCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATT 826
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      136 SerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIle 155

Qy      827 GACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGC 886
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      156 AspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPheGly 175

Qy      887 CTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCA 946
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      176 LeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAla 195

Qy      947 GCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGC 1006
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      196 AlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCys 215

Qy     1007 TGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCAC 1066
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      216 CysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHis 235

Qy     1067 CAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGC 1126
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      236 GlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSer 255
```

```
Qy       1127 TGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTAC 1186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 TrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeuTyr 275

Qy       1187 CAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTT 1246
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 GlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGlyLys 295

Qy       1247 CGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTC 1306
                ::: ||||||           |||   :::        |||   :::::|||   ::
Db        296 AlaLeu-AlaAlaSerArgGlyTrpCysGlySerArgProProGluThrThrLeuGlyAl 315

Qy       1307 ATGGACCGGC 1316
                :   :::|||
Db        315 aValSerGly 318
```

RESULT 21
US-09-780-669-708
; Sequence 708, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780,669
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 708
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-780-669-708

Alignment Scores:

Pred. No.:	2.57e-71	Length:	371
Score:	1420.50	Matches:	275
Percent Similarity:	88.58%	Conservative:	12
Best Local Similarity:	84.88%	Mismatches:	23
Query Match:	22.13%	Indels:	14
DB:	10	Gaps:	1

US-09-759-143-110 (1-3410) x US-09-780-669-708 (1-371)

```
Qy        347 AACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCT 406
              :::|||:::           |||   :::   |||:::       :::|||:::
Db          9 SerLeuValProLeuProLeuAlaLeuTyrLeuSerGlnProLeuThrHisThrThrSer 28

Qy        407 CTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCA 466
              ||||||                                    ||||||||||||
Db         29 LeuLeu------------------------------------AlaGlyIleGlyPro 35

Qy        467 GTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGC 526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         36 ValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArg 55

Qy        527 TATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTT 586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         56 TyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPhe 75

Qy        587 CTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAG 646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 LeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGlu 95

Qy        647 CTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACT 706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 LeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThr 115

Qy        707 CCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTAC 766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        116 ProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyr 135

Qy        767 TCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATT 826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 SerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIle 155

Qy        827 GACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGC 886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 AspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPheGly 175

Qy        887 CTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCA 946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 LeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAla 195

Qy        947 GCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGC 1006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 AlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCys 215

Qy       1007 TGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCAC 1066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 CysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHis 235

Qy       1067 CAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGC 1126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 GlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSer 255

Qy       1127 TGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTAC 1186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 TrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeuTyr 275

Qy       1187 CAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTT 1246
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 GlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGlyLys 295

Qy       1247 CGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTC 1306
                ::: ||||||           |||   :::        |||   :::::|||   ::
Db        296 AlaLeu-AlaAlaSerArgGlyTrpCysGlySerArgProProGluThrThrLeuGlyAl 315

Qy       1307 ATGGACCGGC 1316
                :   :::|||
Db        315 aValSerGly 318
```

RESULT 22
US-09-822-827-708
; Sequence 708, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 708

```
;   LENGTH: 371
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-822-827-708

Alignment Scores:
Pred. No.:               2.57e-71      Length:         371
Score:                   1420.50       Matches:        275
Percent Similarity:      88.58%        Conservative:   12
Best Local Similarity:   84.88%        Mismatches:     23
Query Match:             22.13%        Indels:         14
DB:                      10            Gaps:           1

US-09-759-143-110 (1-3410) x US-09-822-827-708 (1-371)

Qy        347 AACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCT 406
              :::|||:::         |||    :::   |||:::         :::|||:::
Db          9 SerLeuValProLeuProLeuAlaLeuTyrLeuSerGlnProLeuThrHisThrThrSer 28

Qy        407 CTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCA 466
              ||||||                                          ||||||||||||
Db         29 LeuLeu------------------------------------AlaGlyIleGlyPro 35

Qy        467 GTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGC 526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         36 ValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArg 55

Qy        527 TATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTT 586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         56 TyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPhe 75

Qy        587 CTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAG 646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 LeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGlu 95

Qy        647 CTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACT 706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 LeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThr 115

Qy        707 CCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTAC 766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        116 ProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyr 135

Qy        767 TCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATT 826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 SerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIle 155

Qy        827 GACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGC 886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 AspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPheGly 175

Qy        887 CTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCA 946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 LeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAla 195

Qy        947 GCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGC 1006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 AlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCys 215

Qy       1007 TGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCAC 1066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 CysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHis 235

Qy       1067 CAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGC 1126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 GlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSer 255

Qy       1127 TGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTAC 1186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 TrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeuTyr 275

Qy       1187 CAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTT 1246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        276 GlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGlyLys 295

Qy       1247 CGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTC 1306
                 ::: ||||||               |||   :::       |||   :::::|||  ::
Db        296 AlaLeu-AlaAlaSerArgGlyTrpCysGlySerArgProProGluThrThrLeuGlyAl 315

Qy       1307 ATGGACCGGC 1316
               :   :::|||
Db        315 aValSerGly 318

RESULT 23
US-10-012-896-852
; Sequence 852, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Wantanabe, Yoshihiro
;  APPLICANT:  Meagher, Madeleine Joy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C27
;  CURRENT APPLICATION NUMBER: US/10/012,896
;  CURRENT FILING DATE:  2001-12-10
;  NUMBER OF SEQ ID NOS: 1011
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 852
;   LENGTH: 400
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-012-896-852

Alignment Scores:
Pred. No.:               2.25e-70      Length:         400
Score:                   1403.50       Matches:        267
Percent Similarity:      97.45%        Conservative:   0
Best Local Similarity:   97.45%        Mismatches:     2
Query Match:             21.87%        Indels:         5
DB:                      9             Gaps:           1

US-09-759-143-110 (1-3410) x US-10-012-896-852 (1-400)

Qy        371 GTGTGTTTGGCCGCAGGC---------------ATCACCTATGTGCCGCCTCTGCTGCTG 415
                 |||    ||||||   |||               ||||||||||||||||||||||||
Db        127 ValThrLeuAlaGluGlyProProAlaGluPheIleThrTyrValProProLeuLeuLeu 146

Qy        416 GAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGC 475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 GluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGlyProValLeuGly 166

Qy        476 CTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGC 535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 LeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArgTyrGlyArg 186
```

```
Qy        536 CGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCA 595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 ArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPheLeuIlePro 206

Qy        596 AGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTG 655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 ArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGluLeuAlaLeu 226

Qy        656 CTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAG 715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 LeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThrProLeuGlu 246

Qy        716 GCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTAT 775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 AlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyrSerValTyr 266

Qy        776 GCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGAC 835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 AlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIleAspTrpAsp 286

Qy        836 ACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACC 895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPheGlyLeuLeuThr 306

Qy        896 CTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGC 955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 LeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAlaAlaLeuGly 326

Qy        956 CCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGC 1015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 ProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCysCysProCys 346

Qy       1016 CGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGC 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 ArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHisGlnLeuCys 366

Qy       1076 TGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCA 1135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 CysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSerTrpMetAla 386

Qy       1136 CTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAG 1177
              ||||||||||||||||||||||||||||||||||||||||||
Db        387 LeuMetThrPheThrLeuPheTyrThrAspPheValGlyGlu 400
```

RESULT 24

```
US-09-895-793-852
; Sequence 852, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.534C2
;  CURRENT APPLICATION NUMBER: US/09/895,793
;  CURRENT FILING DATE:  2001-06-29
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 852
;   LENGTH: 400
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-895-793-852

Alignment Scores:
Pred. No.:               2.25e-70       Length:        400
Score:                   1403.50        Matches:       267
Percent Similarity:      97.45%         Conservative:  0
Best Local Similarity:   97.45%         Mismatches:    2
Query Match:             21.87%         Indels:        5
DB:                      9              Gaps:          1

US-09-759-143-110 (1-3410) x US-09-895-793-852 (1-400)

Qy        371 GTGTGTTTGGCCGCAGGC---------------ATCACCTATGTGCCGCCTCTGCTGCTG 415
              |||   ||||||   |||               |||||||||||||||||||||||||||
Db        127 ValThrLeuAlaGluGlyProProAlaGluPheIleThrTyrValProProLeuLeuLeu 146

Qy        416 GAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGC 475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 GluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGlyProValLeuGly 166

Qy        476 CTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGC 535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 LeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArgTyrGlyArg 186

Qy        536 CGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCA 595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 ArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPheLeuIlePro 206

Qy        596 AGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTG 655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 ArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGluLeuAlaLeu 226

Qy        656 CTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAG 715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 LeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThrProLeuGlu 246

Qy        716 GCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTAT 775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 AlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyrSerValTyr 266

Qy        776 GCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGAC 835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 AlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIleAspTrpAsp 286

Qy        836 ACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACC 895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPheGlyLeuLeuThr 306

Qy        896 CTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGC 955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 LeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAlaAlaLeuGly 326

Qy        956 CCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGC 1015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 ProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCysCysProCys 346

Qy       1016 CGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGC 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 ArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHisGlnLeuCys 366

Qy       1076 TGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCA 1135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 CysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSerTrpMetAla 386
```

```
Qy       1136 CTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAG 1177
              ||||||||||||||||||||||||||||||||||||||||||
Db        387 LeuMetThrPheThrLeuPheTyrThrAspPheValGlyGlu 400

RESULT 25
US-09-895-814-852
; Sequence 852, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C26
;  CURRENT APPLICATION NUMBER: US/09/895,814
;  CURRENT FILING DATE:  2001-06-29
;  NUMBER OF SEQ ID NOS: 990
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 852
;   LENGTH: 400
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-895-814-852

Alignment Scores:
Pred. No.:                2.25e-70       Length:          400
Score:                    1403.50        Matches:         267
Percent Similarity:       97.45%         Conservative:    0
Best Local Similarity:    97.45%         Mismatches:      2
Query Match:              21.87%         Indels:          5
DB:                       9              Gaps:            1

US-09-759-143-110 (1-3410) x US-09-895-814-852 (1-400)

Qy        371 GTGTGTTTGGCCGCAGGC---------------ATCACCTATGTGCCGCCTCTGCTGCTG 415
              |||   ||||||   |||               |||||||||||||||||||||||||||
Db        127 ValThrLeuAlaGluGlyProProAlaGluPheIleThrTyrValProProLeuLeuLeu 146

Qy        416 GAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGC 475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 GluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGlyProValLeuGly 166

Qy        476 CTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGC 535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 LeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArgTyrGlyArg 186

Qy        536 CGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCA 595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 ArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPheLeuIlePro 206

Qy        596 AGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTG 655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 ArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGluLeuAlaLeu 226

Qy        656 CTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAG 715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 LeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThrProLeuGlu 246

Qy        716 GCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTAT 775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 AlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyrSerValTyr 266

Qy        776 GCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGAC 835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 AlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIleAspTrpAsp 286

Qy        836 ACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACC 895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPheGlyLeuLeuThr 306

Qy        896 CTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGC 955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 LeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAlaAlaLeuGly 326

Qy        956 CCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGC 1015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 ProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCysCysProCys 346

Qy       1016 CGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGC 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 ArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHisGlnLeuCys 366

Qy       1076 TGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCA 1135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 CysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSerTrpMetAla 386

Qy       1136 CTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAG 1177
              ||||||||||||||||||||||||||||||||||||||||||
Db        387 LeuMetThrPheThrLeuPheTyrThrAspPheValGlyGlu 400

RESULT 26
US-09-759-143-852
; Sequence 852, Application US/09759143
; Patent No. US20020022248A1
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li,.Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C23
;  CURRENT APPLICATION NUMBER: US/09/759,143
;  CURRENT FILING DATE:  2001-01-12
;  NUMBER OF SEQ ID NOS:  934
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 852
;   LENGTH: 400
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-759-143-852
```

```
Alignment Scores:
Pred. No.:              2.25e-70       Length:          400
Score:                  1403.50        Matches:         267
Percent Similarity:     97.45%         Conservative:    0
Best Local Similarity:  97.45%         Mismatches:      2
Query Match:            21.87%         Indels:          5
DB:                     10             Gaps:            1

US-09-759-143-110 (1-3410) x US-09-759-143-852 (1-400)

Qy        371 GTGTGTTTGGCCGCAGGC---------------ATCACCTATGTGCCGCCTCTGCTGCTG 415
              |||   ||||||   |||               |||||||||||||||||||||||||
Db        127 ValThrLeuAlaGluGlyProProAlaGluPheIleThrTyrValProProLeuLeuLeu 146

Qy        416 GAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGC 475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 GluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGlyProValLeuGly 166

Qy        476 CTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGC 535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 LeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArgTyrGlyArg 186

Qy        536 CGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCA 595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 ArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPheLeuIlePro 206

Qy        596 AGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTG 655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 ArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGluLeuAlaLeu 226

Qy        656 CTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAG 715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 LeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThrProLeuGlu 246

Qy        716 GCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTAT 775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 AlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyrSerValTyr 266

Qy        776 GCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGAC 835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 AlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIleAspTrpAsp 286

Qy        836 ACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACC 895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPheGlyLeuLeuThr 306

Qy        896 CTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGC 955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 LeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAlaAlaLeuGly 326

Qy        956 CCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGC 1015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 ProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCysCysProCys 346

Qy       1016 CGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGC 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 ArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHisGlnLeuCys 366

Qy       1076 TGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCA 1135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 CysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSerTrpMetAla 386

Qy       1136 CTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAG 1177
              ||||||||||||||||||||||||||||||||||||||||||
Db        387 LeuMetThrPheThrLeuPheTyrThrAspPheValGlyGlu 400
```

```
RESULT 27
US-09-780-669-852
; Sequence 852, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C24
;  CURRENT APPLICATION NUMBER: US/09/780,669
;  CURRENT FILING DATE:  2001-02-09
;  NUMBER OF SEQ ID NOS:  943
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 852
;   LENGTH: 400
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-780-669-852

Alignment Scores:
Pred. No.:              2.25e-70       Length:          400
Score:                  1403.50        Matches:         267
Percent Similarity:     97.45%         Conservative:    0
Best Local Similarity:  97.45%         Mismatches:      2
Query Match:            21.87%         Indels:          5
DB:                     10             Gaps:            1

US-09-759-143-110 (1-3410) x US-09-780-669-852 (1-400)

Qy        371 GTGTGTTTGGCCGCAGGC---------------ATCACCTATGTGCCGCCTCTGCTGCTG 415
              |||   ||||||   |||               |||||||||||||||||||||||||
Db        127 ValThrLeuAlaGluGlyProProAlaGluPheIleThrTyrValProProLeuLeuLeu 146

Qy        416 GAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGC 475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 GluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGlyProValLeuGly 166

Qy        476 CTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGC 535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 LeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArgTyrGlyArg 186

Qy        536 CGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCA 595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 ArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPheLeuIlePro 206

Qy        596 AGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTG 655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 ArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGluLeuAlaLeu 226

Qy        656 CTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAG 715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 LeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThrProLeuGlu 246

Qy        716 GCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTAT 775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 AlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyrSerValTyr 266

Qy        776 GCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGAC 835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 AlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIleAspTrpAsp 286
```

```
Qy        836 ACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACC 895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPheGlyLeuLeuThr 306

Qy        896 CTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGC 955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 LeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAlaAlaLeuGly 326

Qy        956 CCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGC 1015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 ProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCysCysProCys 346

Qy       1016 CGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGC 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 ArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHisGlnLeuCys 366

Qy       1076 TGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCA 1135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 CysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSerTrpMetAla 386

Qy       1136 CTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAG 1177
              ||||||||||||||||||||||||||||||||||||||||||
Db        387 LeuMetThrPheThrLeuPheTyrThrAspPheValGlyGlu 400
```

RESULT 28
US-09-822-827-852
; Sequence 852, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 852
;   LENGTH: 400
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-822-827-852

```
Alignment Scores:
Pred. No.:                2.25e-70     Length:         400
Score:                    1403.50      Matches:        267
Percent Similarity:       97.45%       Conservative:   0
Best Local Similarity:    97.45%       Mismatches:     2
Query Match:              21.87%       Indels:         5
DB:                       10           Gaps:           1

US-09-759-143-110 (1-3410) x US-09-822-827-852 (1-400)

Qy        371 GTGTGTTTGGCCGCAGGC---------------ATCACCTATGTGCCGCCTCTGCTGCTG 415
              |||   ||||||   |||               |||||||||||||||||||||||||||
Db        127 ValThrLeuAlaGluGlyProProAlaGluPheIleThrTyrValProProLeuLeuLeu 146

Qy        416 GAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGC 475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 GluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGlyProValLeuGly 166

Qy        476 CTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGC 535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 LeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArgTyrGlyArg 186

Qy        536 CGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCA 595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 ArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPheLeuIlePro 206

Qy        596 AGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTG 655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 ArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGluLeuAlaLeu 226

Qy        656 CTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAG 715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 LeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThrProLeuGlu 246

Qy        716 GCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTAT 775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 AlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyrSerValTyr 266

Qy        776 GCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGAC 835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 AlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIleAspTrpAsp 286

Qy        836 ACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACC 895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPheGlyLeuLeuThr 306

Qy        896 CTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGC 955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 LeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAlaAlaLeuGly 326

Qy        956 CCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGC 1015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 ProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCysCysProCys 346

Qy       1016 CGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGC 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 ArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHisGlnLeuCys 366

Qy       1076 TGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCA 1135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 CysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSerTrpMetAla 386

Qy       1136 CTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAG 1177
              ||||||||||||||||||||||||||||||||||||||||||
Db        387 LeuMetThrPheThrLeuPheTyrThrAspPheValGlyGlu 400
```

RESULT 29
US-10-012-896-1011
; Sequence 1011, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Wantanabe, Yoshihiro
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012,896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1011

```
;   LENGTH: 355
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-012-896-1011

Alignment Scores:
Pred. No.:               4.66e-45      Length:        355
Score:                   947.00        Matches:       188
Percent Similarity:      99.47%        Conservative:  0
Best Local Similarity:   99.47%        Mismatches:    1
Query Match:             14.76%        Indels:        0
DB:                      9             Gaps:          0

US-09-759-143-110 (1-3410) x US-10-012-896-1011 (1-355)

Qy       1307 ATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCT 1366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 MetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaAla 179

Qy       1367 TTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCA 1426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 PheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAlaSer 199

Qy       1427 GCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCC 1486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 AlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSer 219

Qy       1487 CTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCT 1546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 LeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGlyAla 239

Qy       1547 AGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTC 1606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 SerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaProPhe 259

Qy       1607 CCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGC 1666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 ProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeuCys 279

Qy       1667 GGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGG 1726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAlaArg 299

Qy       1727 GTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTG 1786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 ValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeuLeu 319

Qy       1787 TCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACT 1846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 SerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerValThr 339

Qy       1847 GCCTATATGGTGTCTGCCGCAGGCCTG 1873
              ||||||||||||||||||||    |||
Db        340 AlaTyrMetValSerAlaAlaAlaLeu 348

RESULT 30
US-10-010-940-571
; Sequence 571, Application US/10010940
; Publication No. US20030088062A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan Louise
;  APPLICANT:  Jiang Yuqui
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Kalos, Michael
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Retter, Mark
;  APPLICANT:  Solk, John
;  APPLICANT:  Day, Craig
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427D3
;  CURRENT APPLICATION NUMBER: US/10/010,940
;  CURRENT FILING DATE:  2001-12-05
;  NUMBER OF SEQ ID NOS: 575
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 571
;   LENGTH: 84
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-010-940-571

Alignment Scores:
Pred. No.:               1.37e-17      Length:        84
Score:                   452.00        Matches:       84
Percent Similarity:      100.00%       Conservative:  0
Best Local Similarity:   100.00%       Mismatches:    0
Query Match:             7.04%         Indels:        0
DB:                      9             Gaps:          0

US-09-759-143-110 (1-3410) x US-10-010-940-571 (1-84)

Qy       1493 CACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGT 1552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 HisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGlyAlaSerSer 20

Qy       1553 GAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAAT 1612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaProPheProAsn 40

Qy       1613 GGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCC 1672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 GlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeuCysGlyAla 60

Qy       1673 TCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTT 1732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SerAlaCysAspValSerValArgValValValGlyGluProThrGluAlaArgValVal 80

Qy       1733 CCGGGCCGGGGC 1744
              ||||||||||||
Db         81 ProGlyArgGly 84

RESULT 31
US-10-012-896-706
; Sequence 706, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Wantanabe, Yoshihiro
;  APPLICANT:  Meagher, Madeleine Joy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C27
```

```
;  CURRENT APPLICATION NUMBER: US/10/012,896
;  CURRENT FILING DATE:  2001-12-10
;  NUMBER OF SEQ ID NOS: 1011
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 706
;   LENGTH: 123
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-012-896-706

Alignment Scores:
Pred. No.:                3.74e-16     Length:         123
Score:                    426.00       Matches:        87
Percent Similarity:       97.83%       Conservative:   3
Best Local Similarity:    94.57%       Mismatches:     2
Query Match:              6.64%        Indels:         0
DB:                       9            Gaps:           0

US-09-759-143-110 (1-3410) x US-10-012-896-706 (1-123)

Qy       1250 ATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATG 1309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMet 20

Qy       1310 GACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTC 1369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaAlaPhe 40

Qy       1370 CCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCC 1429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAlaSerAla 60

Qy       1430 GCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTC 1489
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSerLeu 80

Qy       1490 TACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATAC 1525
              ||||||||||||||||||||    :::   ::::::
Db         81 TyrHisArgGluLysGlnValLeuIleGlyGlnTrp 92

RESULT 32
US-09-895-793-706
; Sequence 706, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.534C2
;  CURRENT APPLICATION NUMBER: US/09/895,793
;  CURRENT FILING DATE:  2001-06-29
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 706
;   LENGTH: 123
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-895-793-706

Alignment Scores:
Pred. No.:                3.74e-16     Length:         123
Score:                    426.00       Matches:        87
Percent Similarity:       97.83%       Conservative:   3
Best Local Similarity:    94.57%       Mismatches:     2
Query Match:              6.64%        Indels:         0
DB:                       9            Gaps:           0

US-09-759-143-110 (1-3410) x US-09-895-793-706 (1-123)

Qy       1250 ATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATG 1309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMet 20

Qy       1310 GACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTC 1369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaAlaPhe 40

Qy       1370 CCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCC 1429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAlaSerAla 60

Qy       1430 GCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTC 1489
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSerLeu 80

Qy       1490 TACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATAC 1525
              ||||||||||||||||||||    :::   ::::::
Db         81 TyrHisArgGluLysGlnValLeuIleGlyGlnTrp 92

RESULT 33
US-09-895-814-706
; Sequence 706, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C26
;  CURRENT APPLICATION NUMBER: US/09/895,814
;  CURRENT FILING DATE:  2001-06-29
;  NUMBER OF SEQ ID NOS: 990
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 706
;   LENGTH: 123
```

```
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-895-814-706

Alignment Scores:
Pred. No.:              3.74e-16      Length:        123
Score:                  426.00        Matches:       87
Percent Similarity:     97.83%        Conservative:  3
Best Local Similarity:  94.57%        Mismatches:    2
Query Match:            6.64%         Indels:        0
DB:                     9             Gaps:          0

US-09-759-143-110 (1-3410) x US-09-895-814-706 (1-123)

Qy      1250 ATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATG 1309
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 MetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMet 20

Qy      1310 GACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTC 1369
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        21 AspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaAlaPhe 40

Qy      1370 CCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCC 1429
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        41 ProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAlaSerAla 60

Qy      1430 GCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTC 1489
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 AlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSerLeu 80

Qy      1490 TACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATAC 1525
             |||||||||||||||||||||   :::   ::::::
Db        81 TyrHisArgGluLysGlnValLeuIleGlyGlnTrp 92

RESULT 34
US-09-759-143-706
; Sequence 706, Application US/09759143
; Patent No. US20020022248A1
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C23
;  CURRENT APPLICATION NUMBER: US/09/759,143
;  CURRENT FILING DATE:  2001-01-12
;  NUMBER OF SEQ ID NOS:  934
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 706
;   LENGTH: 123
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-759-143-706

Alignment Scores:
Pred. No.:              3.74e-16      Length:        123
Score:                  426.00        Matches:       87
Percent Similarity:     97.83%        Conservative:  3
Best Local Similarity:  94.57%        Mismatches:    2
Query Match:            6.64%         Indels:        0
DB:                     10            Gaps:          0

US-09-759-143-110 (1-3410) x US-09-759-143-706 (1-123)

Qy      1250 ATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATG 1309
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 MetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMet 20

Qy      1310 GACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTC 1369
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        21 AspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaAlaPhe 40

Qy      1370 CCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCC 1429
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        41 ProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAlaSerAla 60

Qy      1430 GCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTC 1489
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 AlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSerLeu 80

Qy      1490 TACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATAC 1525
             |||||||||||||||||||||   :::   ::::::
Db        81 TyrHisArgGluLysGlnValLeuIleGlyGlnTrp 92

RESULT 35
US-09-780-669-706
; Sequence 706, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C24
;  CURRENT APPLICATION NUMBER: US/09/780,669
;  CURRENT FILING DATE:  2001-02-09
;  NUMBER OF SEQ ID NOS:  943
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 706
;   LENGTH: 123
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-780-669-706

Alignment Scores:
Pred. No.:              3.74e-16      Length:        123
Score:                  426.00        Matches:       87
Percent Similarity:     97.83%        Conservative:  3
Best Local Similarity:  94.57%        Mismatches:    2
Query Match:            6.64%         Indels:        0
DB:                     10            Gaps:          0

US-09-759-143-110 (1-3410) x US-09-780-669-706 (1-123)
```

```
Qy       1250 ATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATG 1309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMet 20

Qy       1310 GACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTC 1369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaAlaPhe 40

Qy       1370 CCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCC 1429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAlaSerAla 60

Qy       1430 GCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTC 1489
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSerLeu 80

Qy       1490 TACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATAC 1525
              ||||||||||||||||||||   :::    ::::::
Db         81 TyrHisArgGluLysGlnValLeuIleGlyGlnTrp 92

RESULT 36
US-09-822-827-706
; Sequence 706, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.534C1
;  CURRENT APPLICATION NUMBER: US/09/822,827
;  CURRENT FILING DATE:  2001-03-28
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 706
;   LENGTH: 123
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-822-827-706

Alignment Scores:
Pred. No.:              3.74e-16      Length:        123
Score:                  426.00        Matches:       87
Percent Similarity:     97.83%        Conservative:  3
Best Local Similarity:  94.57%        Mismatches:    2
Query Match:            6.64%         Indels:        0
DB:                     10            Gaps:          0

US-09-759-143-110 (1-3410) x US-09-822-827-706 (1-123)

Qy       1250 ATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATG 1309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMet 20

Qy       1310 GACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTC 1369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaAlaPhe 40

Qy       1370 CCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCC 1429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAlaSerAla 60

Qy       1430 GCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTC 1489
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSerLeu 80

Qy       1490 TACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATAC 1525
              ||||||||||||||||||||   :::    ::::::
Db         81 TyrHisArgGluLysGlnValLeuIleGlyGlnTrp 92

RESULT 37
US-10-096-986-103
; Sequence 103, Application US/10096986
; Publication No. US20030083464A1
;   GENERAL INFORMATION:
;        APPLICANT: Ferrari, Franco A.
;                   Richardson, Charles
;                   Chambers, James
;                   Causey, Stuart
;                   Pollock, Thomas J.
;                   Cappello, Joseph
;                   Crissman, John W.
;        TITLE OF INVENTION: No. US20030083464A1el Peptides Comprising Repetitive
;                            Units of Amino Acids and DNA Sequences Encoding the Same
;        NUMBER OF SEQUENCES: 117
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Flehr Hohbach Test Albritton & Herbert LLP
;             STREET: Four Embarcadero Center, Suite 3400
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: US
;             ZIP: 94111
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/096,986
;             FILING DATE: 12-Mar-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/444,791
;             FILING DATE: 22-No. US20030083464A1-1999
;             APPLICATION NUMBER: US 08/482,085
;             FILING DATE: 07-JUN-1995
;             APPLICATION NUMBER: US 08/175,155
;             FILING DATE: 29-DEC-1993
;             APPLICATION NUMBER: US 08/053,049
;             FILING DATE: 22-APR-1993
;             APPLICATION NUMBER: US 07/114,618
;             FILING DATE: 29-OCT-1987
;             APPLICATION NUMBER: US 06/927,258
;             FILING DATE: 04-NOV-1986
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Trecartin, Richard F.
;             REGISTRATION NUMBER: 31,801
;             REFERENCE/DOCKET NUMBER: A-55186-11/RFT/BTC
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 415-781-1989
;             TELEFAX: 415-398-3249
;   INFORMATION FOR SEQ ID NO: 103:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 837 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 103:
US-10-096-986-103

Alignment Scores:
Pred. No.:              2.54e-12      Length:        837
Score:                  356.50        Matches:       263
Percent Similarity:     35.16%        Conservative:  32
Best Local Similarity:  31.35%        Mismatches:    327
Query Match:            5.73%         Indels:        218
DB:                     9             Gaps:          57

US-09-759-143-110 (1-3410) x US-10-096-986-103 (1-837)

Qy       2347 GGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCTCAA 2288
                 |||   ||||||                ||||||      |||      |||
Db         46 GlySerProGly---------------AlaProGlyThrPro------GlyProGlnGly 58

Qy       2287 CTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCAGAGTCCCCGCATTCC 2228
```

```
                 |||  |||              |||  |||  |||           |||
Db        59 LeuProGly---------------SerProGlyAlaProGlyThrProGlyProGlnGly 73

Qy      2227 AGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGC 2168
                      ||||||  |||  |||           |||     |||   :::|||
Db        74 LeuProGlySerPro---GlyAlaPro---------GlyThrProGlyPro---GlnGly 88

Qy      2167 CTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCCCCAGCCCCCAGCT 2108
             |||      |||||||||                    |||  |||   |||
Db        89 Leu---ProGlySerPro--------------------GlyAlaProGlyThrProGly 100

Qy      2107 GTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGC 2048
                ||                          |||       |||      |||
Db       101 ProGln------------GlyLeuProGlySerProGlyAlaProGlyThrProGly--P 116

Qy      2047 AACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGGA 1988
               |||   |||      ||||||   ||||||          :::|||
Db       116 roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 136

Qy      1987 CCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGC 1928
                  |||  |||  ||||||
Db       136 lySerProGlyAlaProGlyThrPro---------------------------------- 144

Qy      1927 CAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCC 1868
                         |||  ||||||                         ||||||
Db       145 --------GlyProGlnGlyLeuPro--------------------GlySerProGlyA 155

Qy      1867 TGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCATAAA 1808
                ||||||       |||         |||                  ||||||
Db       155 laProGlyThrProGlyProGlnGlyLeuProGly---------------SerPro---- 168

Qy      1807 CAGGGATGGGGCCACCTGGGACAGCAGGAAGGCAC---------------------TATC 1769
                      |||  |||||||||  |||                           |
Db       169 --------GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaP 186

Qy      1768 CAGGATGGCGAGGTCCAGGCAGATGCCCCGGC------------CCGGAACCACCCTGGC 1721
              |||||     ||||||          ||||||           |||    |||    |
Db       186 roGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyP 206

Qy      1720 CTCGGTG-------GGCTCACCCACCACCACACGTACGGAGACATCACAGGCA------G 1674
              ||:::      |||||||||          |||       |||
Db       206 roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 226

Qy      1673 AGGCCCCGCAGAGCGCGGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTC 1614
                :::|||   :::   |||        |||   :::   ||||||   :::|||     |
Db       226 lySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaP 246

Qy      1613 CATTAGGGAAGGGAGCTCCAGGCTTAGGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCT 1554
              ||        |||     ||||||  ||||||          :::   |||   |||
Db       246 roGlyThrProGlyProGlnGlyLeu---ProGly-SerProGlyAlaProGlyThrPro 264

Qy      1553 CACTGCTAGCACCTCCAGTGTCCCCTCGGTATT------TGGGCAGGAACACCTGCTTCT 1500
                                          |||   |||:::         |||   |||   :::
Db       265 Gly------------------ProGlnGlyLeuProGlySerProGlyAlaProGlyThr 278

Qy      1499 CCCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACC 1440
             ||||||   :::|||   |||             ||||||   |||        |||   ||
Db       279 ProGlyProGlnGlyLeuPro------------GlySerProGlyAla-ProGlyThrPr 294

Qy      1439 CGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACC-- 1382
              ||||     |||                |||                |||   ||||||
Db       294 oGlyProGlnGlyLeuProGlySerProGly---------------AlaProGlyThrPr 309

Qy      1381 -GGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCT 1323
                |||  |||||||||   ||||||       |||
Db       309 oGly--ProGlnGly---LeuProGlySerPro-----------------------Gly 319

Qy      1322 GCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACA--- 1266
             ||||||       |||   |||:::         ||||||               ||||||
Db       320 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 339

Qy      1265 ---------------------------GCCCCAGGCTGCCCATCC---GAACGCCTT 1239
                                        |||||||||    |||         :::     |||
Db       340 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 359

Qy      1238 CATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTACAGCC 1179
                  |||        |||       |||                  ||||||:::
Db       360 ProGlySerProGlyAlaProGlyThrProGlyPro---GlnGlyLeuProGlySerPro 378

Qy      1178 CCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGC------ 1125
                  |||      |||                              :::|||   |||
Db       379 GlyAlaProGlyThrProGlyProGlnGly-----------LeuProGlySerProGly 394

Qy      1124 ------TGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACA 1071
                   |||       |||:::               ||||||   |||:::
Db       395 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 414

Qy      1070 GCT-----GGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGG--- 1020
                      |||     |||      |||   ||||||:::         |||   |||
Db       415 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 434

Qy      1019 CCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGG 960
             ||||||         :::      |||      ||||||      |||         :::
Db       435 ProGlySerProGlyAlaProGlyThr---ProGlyProGlnGlyLeuProGlySerPro 453

Qy       959 TGGGGCCCAGCGCTGCCT---CCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGA 903
                  |||     |||   ||||||                                 |||
Db       454 GlyAlaProGlyThrProGlyProGln-----------------------------Gly 463

Qy       902 AGATGAGGGTGAGCAGGCCAAAGAGGCACTCCT---------------CCTGGGTGCCCA 858
                  |||      |||    |||||||||                  ||||||   |||
Db       464 LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro 483

Qy       857 GGTAGGGGGCCAGGGCACTGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGC 798
             |||       ||||||          |||                  |||||||||:::
Db       484 Gly---AlaProGly------ThrProGlyProGlnGlyLeuProGlySerProGlyAla 500

Qy       797 CCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGT 741
             |||                               ||||||   |||   ||||||
Db       501 Pro-------------------------GlyThrProGlyProGlnGlyLeuProGly 511

Qy       740 ---CCCGGAAGAGGTCAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGA 684
                ||||||                ||||||            :::||||||
Db       512 SerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGly------ 529

Qy       683 AGTCCAGCAGCCCCACGCCCAGGATGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGC 624
              :::|||      |||   |||                :::|||||||||      ||||||
Db       530 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla------ProGly 547

Qy       623 ACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCA 564
             |||                |||   |||         ||||||   |||   |||
Db       548 ThrProGly--------------ProGlnGly---LeuProGlySerProGlyAlaPro 561

Qy       563 AGGACAGTG---CCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCAC 507
                 |||     |||:::          |||
Db       562 GlyThrProGlyProGlnGlyLeuProGly---------------------------- 571

Qy       506 TGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCA 447
                   |||  ||||||        |||||||||                        |||
Db       572 ------SerProGlyAlaProGlyThrProGlyProGln-----------GlyLeuPro 585

Qy       446 TGGTCATGAACTTCTCCTCTACCCCCA-----------CTTCCAGCAGCAGAGGCGGCA 399
                |||      :::|||   |||              ||||||   :::      |||
Db       586 GlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 605

Qy       398 CATAGGTGATGCCTGCGGCCAAACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCA 339
                                   |||   ||||||   |||   |||   |||   |||
Db       606 ------------------ProGlyThrProGlyProGlnGlyLeuProGlySerProGly 619

Qy       338 AGAGCTGGGCTTTCCGGT---GCCGCAGCAGGCGGCTCACCCACAGCCTCTGGACCATAG 282
                  |||     |||              |||||||||   |||   |||
Db       620 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr-Pr 639

Qy       281 TGGGCCAGGCGGG---TAGGGCTCAGGGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCT 225
              ||||||   |||      ||||||   |||       |||||||||              ||
Db       639 oGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLe 659
```

```
Qy     224 CGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAACTGCCTAGGAATCA 165
           |       :::|||   |||        :::|||                   |||:::
Db     659 uProGlySerProGlyAlaProGlyThrProGly---------------ProGlnGlyLe 674

Qy     164 GCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGCTTCTCAGCCCATGCTCA- 106
             |||||||:::                |||       ||||||      |||   |||
Db     674 uProGlySerProGlyAla---ProGlyThrProGlyProGlnGly-LeuProGlySerP 693

Qy     105 --------------ACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTCATCACT 60
                         ||||||      |||    |||    |||
Db     693 roGlyAlaProGlyThrProGlyProGlnGly---LeuProGlySerProGlyAlaProG 712

Qy      59 CAGATCCTGGC---CGAGGCGCGCGGCTGTCACCCGGAGCC 22
                  ||||||   :::|||     .   ||||||||||||
Db     712 lyThrProGlyProGlnGlyLeuProGlySerProGlyAla 725
```

RESULT 38
US-10-216-705-21
; Sequence 21, Application US/10216705
; Publication No. US20030096973A1
; GENERAL INFORMATION:
; APPLICANT: Meristem Therapeutics, S.A.
; TITLE OF INVENTION: Recombinant Collagens and Derived Proteins Produced by Plants, Me
; TITLE OF INVENTION: obtaining Such and Their Uses
; FILE REFERENCE: 1149-3 DIV
; CURRENT APPLICATION NUMBER: US/10/216,705
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 09/331,347
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
;   LENGTH: 1464
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-216-705-21

```
Alignment Scores:
Pred. No.:               8.41e-12      Length:          1464
Score:                   347.00        Matches:         267
Percent Similarity:      32.42%        Conservative:    42
Best Local Similarity:   28.02%        Mismatches:      330
Query Match:             5.57%         Indels:          315
DB:                      9             Gaps:            55
```

US-09-759-143-110 (1-3410) x US-10-216-705-21 (1-1464)

```
Qy    2353 AGAGATGGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCT------TCTCTA 2300
           |||||||||   ||||||                   |||       |||
Db     131 ArgAspGlyIleProGlyGlnProGlyLeuProGlyProProGlyProProGlyProPro 150

Qy    2299 GGTGTGTCTCAACTAGGAGGCTAGCTGTTA---ACCCTGAGCCTGGGTAATCCACCTGCA 2243
            |||             |||||||||  |||          ||||||   |||
Db     151 GlyProProGlyLeuGlyGlyLysLeuAlaProGlnLeuSerTyrGlyTyrAspGluLys 170

Qy    2242 GAGTCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAA 2183
             :::         |||         |||   |||
Db     171 SerThrGlyGlyIleSerValProGlyProMetGly------------------------ 182

Qy    2182 ACCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACG 2123
                                ||||||                   ||||||
Db     183 ------------------ProSerGlyPro--------------------ArgGlyLeu 189

Qy    2122 CCCCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGAGCCA 2063
            |||    ||||||                          ||||||       ||||||
Db     190 ProGlyProProGlyAlaPro---GlyProGlnGlyPheGlnGlyProProGlyGluPro 208

Qy    2062 CATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAACAGG 2003
                    |||::::::              |||        |||          :::
Db     209 GlyGluProGlyAlaSerGlyProMetGlyProArgGlyProProGlyProProGlyLys 228

Qy    2002 AGCGGGGAGCTGGGA--------------------CCCAGTGAGGCAGGCCCTCCA--- 1967
              :::|||:::   |||                 |||    |||    |||||||||
Db     229 AsnGlyAspAspGlyGluAlaGlyLysProGlyArgProGlyGluArgGlyProProGly 248

Qy    1966 ---------------------CCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGC 1928
                                |||    ||||||               :::    |||
Db     249 ProGlnGlyAlaArgGlyLeuProGlyThrAlaGly---------------LeuProGly 263

Qy    1927 CAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCC 1868
                             :::                         ||||||      |||
Db     264 MetLysGlyHisArgGlyPheSerGlyLeuAspGlyAlaLysGlyAspAlaGlyProAla 283

Qy    1867 TGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCATAAA 1808
              :::         ||||||                  :::||||||||||
Db     284 GlyProLysGlyGluProGlySerProGly------------GluAsnGlyAlaProGly 299

Qy    1807 CAGGGATGGGGCCA---------CCTGGGACAGCAGGAAGGCACTATCCAGGATGGCGAG 1757
           |||  |||||||||         ||||||      ||||||       ||||||     |
Db     300 Gln--MetGlyProArgGlyLeuProGlyGluArgGlyArg------ProGlyAlaProG 317

Qy    1756 GTCCAGGCAGATGCCCCGGCCCGGAA--------------------CCACCCTGGCCTC 1718
           |||||         |||     :::                         ||||||   |||
Db     317 lyProAlaGlyAlaArgGlyAsnAspGlyAlaThrGlyAlaAlaGlyProProGlyProT 337

Qy    1717 GGTGGGCTCACCCACCACCACA---------------------CGTACGGAGACATCAC 1680
                     ||||||                                 :::   |||
Db     337 hrGlyProAlaGlyProProGlyPheProGlyAlaValGlyAlaLysGlyGluAlaGlyP 357

Qy    1679 AG---------------GCAGAGGCCCCGCAGAGCGCGGGTGGAGGTGGGAGCAGGCCAC 1635
                            :::|||  ||||||        |||              |||
Db     357 roGlnGlyProArgGlySerGluGlyProGlnGlyValArgGly---------GluProG 374

Qy    1634 TGCCTCCAGCACCCACGTGTCCATTAGGG---AAGGGAGCTCCAGGCTTAGGG------C 1584
              ||||||   |||                |||    |||   ||||||          |
Db     374 lyProProGlyProAlaGlyAlaAlaGlyProAlaGlyAsnProGlyAlaAspGlyGlnP 394

Qy    1583 CTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGT 1524
            |||||   |||      ::::::|||   |||                          |||
Db     394 roGlyAlaLys-GlyAlaAsnGlyAlaPro--------------------------Gly 404

Qy    1523 ATTTGGGCAGGAACACCTGCTTCTCCCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCAGGA 1464
            |||   ||||||  |||       ||||||    ||||||            :::|||
Db     405 Ile---AlaGlyAlaProGlyPheProGlyAlaArgGly------------ProSerGly 419

Qy    1463 TCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACAC 1404
              ||||||      |||    ||||||     |||   |||          |||
Db     420 Pro-GlnGlyProGlyGlyProProGlyProLysGlyAsnSerGlyGluProGly---Al 438

Qy    1403 TGTGGGACAGGCATGTGGCACCGGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGA 1344
             |||              |||||| |||   |||:::
Db     438 aProGlySerLysGlyAspThrGly-AlaLysGlyGluProGlyPro------------ 453

Qy    1343 CTGCTCGAGTGCCGAATCGCTGCACCAGCCGGTCCATGACCAGAG------AGAAGACCA 1290
                     |||   :::      |||   ||||||      |||       :::|||
Db     454 --------Val-GlyValGlnGlyProProGlyProAlaGlyGluGluGlyLysArgGly 470

Qy    1289 GGGAGATGGCGCACTGCAGGAACAGCC---CCAGGCTGCCCATCCGAACGC--------- 1242
              |||           |||      ||||||   |||    ||||||
Db     471 AlaArgGlyGluProGlyProThrGlyLeuProGlyProProGlyGluArgGlyGlyPro 490

Qy    1241 ------------------CTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTC 1200
                                          |||:::|||||||:::   ||||||
Db     491 GlySerArgGlyPheProGlyAlaAspGlyValAlaGlyProLysGlyProAlaGlyGlu 510

Qy    1199 TGGGCACGCCCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCA 1140
                 ||||||   |||                :::|||
Db     511 ArgGlySerProGlyProAlaGlyProLysGlySerPro--------------------- 523

Qy    1139 TGAGTGCCATCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGC 1080

Db     523 ------------------------------------------------------------ 523

Qy    1079 GGCAGCACACAGCTGGTGCAGCCGGGGAAGCAGGGCGCCCAGGT------------TCCGGA 1032
                 ||||||       ||||||||||||   ||||||                :::|||
```

```
Db      524 ---GlyGluAlaGlyArgProGlyGluAlaGlyLeuProGlyAlaLysGlyLeuThrGly 542

Qy     1031 AAGCCAAGCGGGCCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTT 972
               |||      ||||||  |||                :::  ||||||
Db      543 SerProGlySerProGlyProAspGly------------LysThrGlyProPro------ 556

Qy      971 CTGCTGGCTCGGTGGGGCCCAGCG------------------CTGCCTCCTCAGCCACCA 930
                                |||||||||           ||||||  ||||||
Db      557 ---------------GlyProAlaGlyGlnAspGlyArgProGlyProProGlyProPro 571

Qy      929 ------------------------------------------------GCAGTGTGGCTG 918
                                                              |||
Db      572 GlyAlaArgGlyGlnAlaGlyValMetGlyPheProGlyProLysGlyAlaAla------ 589

Qy      917 CTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGCACT------CCTCCTGGG 864
                        |||      |||  ||||||:::||||||         |||||||||
Db      590 ------------GlyGluProGlyLysAlaGlyGluArgGlyValProGlyProProGly 605

Qy      863 TGCCCAGGT-----------------------------------------------AGG 852
                  |||
Db      606 AlaValGlyProAlaGlyLysAspGlyGluAlaGlyAlaGlnGlyProProGlyProAla 625

Qy      851 GGGCCAGGGCACTGGTGTCCCAGTCAATGGCAGGCAGGA-------------------- 813
            ||||||              |||      ||||||
Db      626 GlyProAlaGlyGluArgGlyGluGlnGlyProAlaGlySerProGlyPheGlnGlyLeu 645

Qy      812 ---GGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGC 756
               |||      |||   |||    :::                 |||      ||||||
Db      646 ProGlyProAlaGlyProProGlyGluAlaGlyLysProGlyGluGlnGlyValProGly 665

Qy      755 GACAGTGGT---CCGGGTCCC------GGAAGAGGTCAGAGAGCAGGGCCTCCAGTGGAG 705
            |||    |||  |||||||||       ||||||:::|||    |||   |||   |||
Db      666 AspLeuGlyAlaProGlyProSerGlyAlaArgGlyGluArg---GlyPheProGlyGlu 684

Qy      704 TGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGCCCAGGATGAGCAGTGCCA--- 648
                     |||         ||||||  ||||||                   |||
Db      685 ArgGlyValGlnGlyProProGlyProAlaGlyProArgGlyAlaAsnGlyAlaProGly 704

Qy      647 ---------------------------GCTCCAGGGGCCTGGGATCCGGGCACAGCAGCC 615
                                       ||||||||||||      ||||||:::
Db      705 AsnAspGlyAlaLysGlyAspAlaGlyAlaProGlyAla------ProGlySerGlnGly 722

Qy      614 CTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCAAGGACAGTG 555
                   :::    :::         ||||||||:::||||||    |||
Db      723 AlaProGlyLeuGlnGlyMetProGlyGluArgGlyAlaAlaGlyLeuProGly------ 740

Qy      554 CCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTA 495
            |||:::       |||                  |||   |||        |||
Db      741 ProLysGlyAspArgGlyAspAlaGlyProLysGlyAlaAspGly---------SerPro 757

Qy      494 GGAGCG---GGACACAGA-----------------------CCAGGCCCAGCACTGGAC 462
            |||      |||   |||                        ||||||||||||
Db      758 GlyLysAspGlyValArgGlyLeuThrGlyProIleGlyProProGlyProAla------ 775

Qy      461 CAATGCCCA-------------------------------------------------- 453
                  |||
Db      776 GlyAlaProGlyAspLysGlyGluSerGlyProSerGlyProAlaGlyProThrGlyAla 795

Qy      452 ------GCACCATGGTCATGAACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCA 399
                  ||||||            |||   ||||||   ||||||        |||
Db      796 ArgGlyAlaProGlyAspArgGlyGluProGlyProProGlyProAlaGlyPheAlaGly 815

Qy      398 CATAGGTGATGCCTGCGGCCAAACACACCTCCAGGC------------CAAAGGTTAGCA 351
                                       |||||||||
Db      816 ---------------------------ProProGlyAlaAspGlyGlnProGlyAlaLys 826

Qy      350 GGTTGACCAGCAAGAGCTGGGCTTTCCGGTGCCGCAGCAGGCGGCTCACCCACAGCCTCT 291
            |||   |||      ||||||     |||   |||         |||   |||      :::
Db      827 GlyGluProGlyAspAlaGlyAlaLysGlyAspAla------GlyProProGlyProAla 844

Qy      290 GGACCATAGTGGGCCA------------GGCGGGTAGGGCTCAGGGGGCCGTTCAGGCAC 243
            |||||| :::||||||             |||       |||:::   |||      |||
Db      845 GlyPro-AlaGlyProProGlyProIleGlyAsnValGlyAlaProGlyAlaLysGlyAl 864

Qy      242 TCCAGAACTGCTTCGTCTCGGCTCTGCT------CCAGAAGCTGCGGCCTCTCCTCCTTG 189
                                       |||||||||    |||  |||        |||
Db      864 a---------------ArgGlySerAlaGlyProProGlyAlaThrGlyPheProGlyAl 879

Qy      188 CTGCCGCCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGG 129
                 ::::::                |||      :::         |||         ||
Db      879 aAlaGlyArgValGlyProProGlyProSerGlyAsnAlaGlyProProGlyProProGl 899

Qy      128 TCCAGCTTCTCAGCCCATGCTCAACACCTGCTGCTGTGGG------GCACCTCAGTGGGG 75
             |||||||    :::                         |||         |||:::  |||:
Db      899 yProAlaGlyLysGlu--------------------GlyGlyLysGlyProArg-GlyG 912

Qy       74 ACACGTCTCATCACTCAGATCCTGGCCGA---------------------------G 45
             :::|||            ||||||                                |
Db      912 luThrGlyProAlaGlyArgProGlyGluValGlyProProGlyProProGlyProAlaG 932

Qy       44 GCGCGCGGCTGTCACCCGGAGCC 22
             ||   :::   ||||||||||||
Db      932 lyGluLysGlySerProGlyAla 939
```

RESULT 39
US-10-060-036-159
; Sequence 159, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yuqiu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159
;   LENGTH: 1464
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-060-036-159

Alignment Scores:

Pred. No.:	1.4e-11	Length:	1464
Score:	343.00	Matches:	266
Percent Similarity:	32.32%	Conservative:	42
Best Local Similarity:	27.91%	Mismatches:	331
Query Match:	5.51%	Indels:	315
DB:	9	Gaps:	55

US-09-759-143-110 (1-3410) x US-10-060-036-159 (1-1464)

```
Qy     2353 AGAGATGGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCT------TCTCTA 2300
             |||||||||  ||||||                  |||      |||
Db      131 ArgAspGlyIleProGlyGlnProGlyLeuProGlyProProGlyProProGlyProPro 150

Qy     2299 GGTGTGTCTCAACTAGGAGGCTAGCTGTTA---ACCCTGAGCCTGGGTAATCCACCTGCA 2243
             |||       |||||||||                   ||||||    |||
Db      151 GlyProProGlyLeuGlyGlyAsnPheAlaProGlnLeuSerTyrGlyTyrAspGluLys 170

Qy     2242 GAGTCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAA 2183
                :::          |||         |||   |||
Db      171 SerThrGlyGlyIleSerValProGlyProMetGly------------------------ 182

Qy     2182 ACCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACG 2123
                                ||||||                             ||||||
Db      183 ------------------ProSerGlyPro--------------------ArgGlyLeu 189

Qy     2122 CCCCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGAGCCA 2063
             |||    ||||||                                ||||||       ||||||
```

```
Db    190 ProGlyProProGlyAlaPro---GlyProGlnGlyPheGlnGlyProProGlyGluPro 208

Qy   2062 CATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAACAGG 2003

Db    209 GlyGluProGlyAlaSerGlyProMetGlyProArgGlyProProGlyProProGlyLys 228

Qy   2002 AGCGGGGAGCTGGGA--------------------CCCAGTGAGGCAGGCCCTCCA--- 1967

Db    229 AsnGlyAspAspGlyGluAlaGlyLysProGlyArgProGlyGluArgGlyProProGly 248

Qy   1966 --------------------CCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGC 1928

Db    249 ProGlnGlyAlaArgGlyLeuProGlyThrAlaGly--------------LeuProGly 263

Qy   1927 CAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCC 1868

Db    264 MetLysGlyHisArgGlyPheSerGlyLeuAspGlyAlaLysGlyAspAlaGlyProAla 283

Qy   1867 TGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCATAAA 1808

Db    284 GlyProLysGlyGluProGlySerProGly-----------GluAsnGlyAlaProGly 299

Qy   1807 CAGGGATGGGGCCA--------CCTGGGACAGCAGGAAGGCACTATCCAGGATGGCGAG 1757

Db    300 Gln--MetGlyProArgGlyLeuProGlyGluArgGlyArg------ProGlyAlaProG 317

Qy   1756 GTCCAGGCAGATGCCCCGGCCCGGAA--------------------CCACCCTGGCCTC 1718

Db    317 lyProAlaGlyAlaArgGlyAsnAspGlyAlaThrGlyAlaAlaGlyProProGlyProT 337

Qy   1717 GGTGGGCTCACCCACCACCACA--------------------CGTACGGAGACATCAC 1680

Db    337 hrGlyProAlaGlyProProGlyPheProGlyAlaValGlyAlaLysGlyGluAlaGlyP 357

Qy   1679 AG--------------GCAGAGGCCCCGCAGAGCGCGGGTGGAGGTGGGAGCAGGCCAC 1635

Db    357 roGlnGlyProArgGlySerGluGlyProGlnGlyValArgGly---------GluProG 374

Qy   1634 TGCCTCCAGCACCCACGTGTCCATTAGGG---AAGGGAGCTCCAGGCTTAGGG------C 1584

Db    374 lyProProGlyProAlaGlyAlaAlaGlyProAlaGlyAsnProGlyAlaAspGlyGlnP 394

Qy   1583 CTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGT 1524

Db    394 roGlyAlaLys-GlyAlaAsnGlyAlaPro-------------------------Gly 404

Qy   1523 ATTTGGGCAGGAACACCTGCTTCTCCCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCAGGA 1464

Db    405 Ile---AlaGlyAlaProGlyPheProGlyAlaArgGly-----------ProSerGly 419

Qy   1463 TCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACAC 1404

Db    420 Pro-GlnGlyProGlyGlyProProGlyProLysGlyAsnSerGlyGluProGly---Al 438

Qy   1403 TGTGGGACAGGCATGTGGCACCGGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGA 1344

Db    438 aProGlySerLysGlyAspThrGly-AlaLysGlyGluProGlyPro------------ 453

Qy   1343 CTGCTCGAGTGCCGAATCGCTGCACCAGCCGGTCCATGACCAGAG------AGAAGACCA 1290

Db    454 --------Val-GlyValGlnGlyProProGlyProAlaGlyGluGluGlyLysArgGly 470

Qy   1289 GGGAGATGGCGCACTGCAGGAACAGCC---CCAGGCTGCCCATCCGAACGC--------- 1242

Db    471 AlaArgGlyGluProGlyProThrGlyLeuProGlyProProGlyGluArgGlyGlyPro 490

Qy   1241 -----------------CTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTC 1200

Db    491 GlySerArgGlyPheProGlyAlaAspGlyValAlaGlyProLysGlyProAlaGlyGlu 510

Qy   1199 TGGGCACGCCCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCA 1140

Db    511 ArgGlySerProGlyProAlaGlyProLysGlySerPro-------------------- 523

Qy   1139 TGAGTGCCATCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGC 1080

Db    523 ------------------------------------------------------------ 523

Qy   1079 GGCAGCACAGCTGGTGCAGCCGGGGAAGCAGGGCGCCCAGGT-----------TCCGGA 1032

Db    524 ---GlyGluAlaGlyArgProGlyGluAlaGlyLeuProGlyAlaLysGlyLeuThrGly 542

Qy   1031 AAGCCAAGCGGGCCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTT 972

Db    543 SerProGlySerProGlyProAspGly------------LysThrGlyProPro------ 556

Qy    971 CTGCTGGCTCGGTGGGGCCCAGCG-----------------CTGCCTCCTCAGCCACCA 930

Db    557 --------------GlyProAlaGlyGlnAspGlyArgProGlyProProGlyProPro 571

Qy    929 ---------------------------------------------GCAGTGTGGCTG 918

Db    572 GlyAlaArgGlyGlnAlaGlyValMetGlyPheProGlyProLysGlyAlaAla------ 589

Qy    917 CTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGCACT------CCTCCTGGG 864

Db    590 -----------GlyGluProGlyLysAlaGlyGluArgGlyValProGlyProProGly 605

Qy    863 TGCCCAGGT-------------------------------------------AGG 852

Db    606 AlaValGlyProAlaGlyLysAspGlyGluAlaGlyAlaGlnGlyProProGlyProAla 625

Qy    851 GGGCCAGGGCACTGGTGTCCCAGTCAATGGCAGGCAGGA-------------------- 813

Db    626 GlyProAlaGlyGluArgGlyGluGlnGlyProAlaGlySerProGlyPheGlnGlyLeu 645

Qy    812 ---GGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGC 756

Db    646 ProGlyProAlaGlyProProGlyGluAlaGlyLysProGlyGluGlnGlyValProGly 665

Qy    755 GACAGTGGT---CCGGGTCCC------GGAAGAGGTCAGAGAGCAGGGCCTCCAGTGGAG 705

Db    666 AspLeuGlyAlaProGlyProSerGlyAlaArgGlyGluArg---GlyPheProGlyGlu 684

Qy    704 TGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGCCCAGGATGAGCAGTGCCA--- 648

Db    685 ArgGlyValGlnGlyProProGlyProAlaGlyProArgGlyAlaAsnGlyAlaProGly 704

Qy    647 -------------------------GCTCCAGGGGCCTGGGATCCGGGCACAGCAGCC 615

Db    705 AsnAspGlyAlaLysGlyAspAlaGlyAlaProGlyAla------ProGlySerGlnGly 722

Qy    614 CTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCAAGGACAGTG 555

Db    723 AlaProGlyLeuGlnGlyMetProGlyGluArgGlyAlaAlaGlyLeuProGly------ 740

Qy    554 CCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTA 495

Db    741 ProLysGlyAspArgGlyAspAlaGlyProLysGlyAlaAspGly---------SerPro 757

Qy    494 GGAGCG---GGACACAGA----------------------CCAGGCCCAGCACTGGAC 462

Db    758 GlyLysAspGlyValArgGlyLeuThrGlyProIleGlyProProGlyProAla------ 775

Qy    461 CAATGCCCA--------------------------------------------------- 453

Db    776 GlyAlaProGlyAspLysGlyGluSerGlyProSerGlyProAlaGlyProThrGlyAla 795

Qy    452 ------GCACCATGGTCATGAACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCA 399

Db    796 ArgGlyAlaProGlyAspArgGlyGluProGlyProProGlyProAlaGlyPheAlaGly 815

Qy    398 CATAGGTGATGCCTGCGGCCAAACACACCTCCAGGC-----------CAAAGGTTAGCA 351

Db    816 -------------------------ProProGlyAlaAspGlyGlnProGlyAlaLys 826
```

```
Qy        350 GGTTGACCAGCAAGAGCTGGGCTTTCCGGTGCCGCAGCAGGCGGCTCACCCACAGCCTCT 291
              |||  |||     ||||||     |||  |||     |||  |||     :::
Db        827 GlyGluProGlyAspAlaGlyAlaLysGlyAspAla------GlyProProGlyProAla 844

Qy        290 GGACCATAGTGGGCCA-----------GGCGGGTAGGGCTCAGGGGGCCGTTCAGGCAC 243
              |||||| :::||||||           |||     |||:::  |||     |||
Db        845 GlyPro-AlaGlyProProGlyProIleGlyAsnValGlyAlaProGlyAlaLysGlyAl 864

Qy        242 TCCAGAACTGCTTCGTCTCGGCTCTGCT------CCAGAAGCTGCGGCCTCTCCTCCTTG 189
                               |||||||||      |||   |||        |||
Db        864 a--------------ArgGlySerAlaGlyProProGlyAlaThrGlyPheProGlyAl 879

Qy        188 CTGCCGCCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGG 129
                      ::::::          |||      :::        |||       ||
Db        879 aAlaGlyArgValGlyProProGlyProSerGlyAsnAlaGlyProProGlyProProGl 899

Qy        128 TCCAGCTTCTCAGCCCATGCTCAACACCTGCTGCTGTGGG------GCACCTCAGTGGGG 75
              |||||||   :::                           |||       |||::: |||:
Db        899 yProAlaGlyLysGlu--------------------GlyGlyLysGlyProArg-GlyG 912

Qy         74 ACACGTCTCATCACTCAGATCCTGGCCGA----------------------------G 45
              ::|||           ||||||                                  |
Db        912 luThrGlyProAlaGlyArgProGlyGluValGlyProProGlyProProGlyProAlaG 932

Qy         44 GCGCGCGGCTGTCACCCGGAGCC 22
              ||    :::   |||||||||||
Db        932 lyGluLysGlySerProGlyAla 939
```

RESULT 40
US-10-171-311-36
; Sequence 36, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersh, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 1464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-36

Alignment Scores:

Pred. No.:	1.4e-11	Length:	1464
Score:	343.00	Matches:	266
Percent Similarity:	32.32%	Conservative:	42
Best Local Similarity:	27.91%	Mismatches:	331
Query Match:	5.51%	Indels:	315
DB:	9	Gaps:	55

US-09-759-143-110 (1-3410) x US-10-171-311-36 (1-1464)

```
Qy       2353 AGAGATGGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCT------TCTCTA 2300
              |||||||||   ||||||                    |||       |||
Db        131 ArgAspGlyIleProGlyGlnProGlyLeuProGlyProProGlyProProGlyProPro 150

Qy       2299 GGTGTGTCTCAACTAGGAGGCTAGCTGTTA---ACCCTGAGCCTGGGTAATCCACCTGCA 2243
              |||         |||||||||                 ||||||   |||
Db        151 GlyProProGlyLeuGlyGlyAsnPheAlaProGlnLeuSerTyrGlyTyrAspGluLys 170

Qy       2242 GAGTCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAA 2183
                :::           |||         |||   |||
Db        171 SerThrGlyGlyIleSerValProGlyProMetGly----------------------- 182

Qy       2182 ACCCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACG 2123
                                 ||||||                          ||||||
Db        183 -----------------ProSerGlyPro--------------------ArgGlyLeu 189

Qy       2122 CCCCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGAGCCA 2063
              |.||    ||||||                      ||||||          ||||||
Db        190 ProGlyProProGlyAlaPro---GlyProGlnGlyPheGlnGlyProProGlyGluPro 208

Qy       2062 CATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAACAGG 2003
                       |||::::::           |||          |||          :::
Db        209 GlyGluProGlyAlaSerGlyProMetGlyProArgGlyProProGlyProProGlyLys 228

Qy       2002 AGCGGGGAGCTGGGA-------------------CCCAGTGAGGCAGGCCCTCCA--- 1967
              :::|||:::   |||                   |||   |||   ||||||||
Db        229 AsnGlyAspAspGlyGluAlaGlyLysProGlyArgProGlyGluArgGlyProProGly 248

Qy       1966 --------------------CCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGC 1928
                                  |||     ||||||               :::   |||
Db        249 ProGlnGlyAlaArgGlyLeuProGlyThrAlaGly--------------LeuProGly 263

Qy       1927 CAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCC 1868
                                 :::                 ||||||         |||
Db        264 MetLysGlyHisArgGlyPheSerGlyLeuAspGlyAlaLysGlyAspAlaGlyProAla 283

Qy       1867 TGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCATAAA 1808
                 :::       ||||||                     :::|||||||||
Db        284 GlyProLysGlyGluProGlySerProGly-----------GluAsnGlyAlaProGly 299

Qy       1807 CAGGGATGGGGCCA---------CCTGGGACAGCAGGAAGGCACTATCCAGGATGGCGAG 1757
              |||  |||||||||         ||||||     ||||||      ||||||     |
Db        300 Gln--MetGlyProArgGlyLeuProGlyGluArgGlyArg------ProGlyAlaProG 317

Qy       1756 GTCCAGGCAGATGCCCCGGCCCGGAA--------------------CCACCCTGGCCTC 1718
              |||||          |||     :::                   ||||||   |||
Db        317 lyProAlaGlyAlaArgGlyAsnAspGlyAlaThrGlyAlaAlaGlyProProGlyProT 337

Qy       1717 GGTGGGCTCACCCACCACCACA---------------------CGTACGGAGACATCAC 1680
                            ||||||                      :::   |||
Db        337 hrGlyProAlaGlyProProGlyPheProGlyAlaValGlyAlaLysGlyGluAlaGlyP 357

Qy       1679 AG--------------GCAGAGGCCCCGCAGAGCGCGGGTGGAGGTGGGAGCAGGCCAC 1635
                              :::|||   ||||||         |||           |||
Db        357 roGlnGlyProArgGlySerGluGlyProGlnGlyValArgGly---------GluProG 374

Qy       1634 TGCCTCCAGCACCCACGTGTCCATTAGGG---AAGGGAGCTCCAGGCTTAGGG------C 1584
                ||||||   |||             |||      |||    |||||         |
Db        374 lyProProGlyProAlaGlyAlaAlaGlyProAlaGlyAsnProGlyAlaAspGlyGlnP 394

Qy       1583 CTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGT 1524
              |||||   |||     ::::::|||   |||                          |||
Db        394 roGlyAlaLys-GlyAlaAsnGlyAlaPro-------------------------Gly 404

Qy       1523 ATTTGGGCAGGAACACCTGCTTCTCCCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCAGGA 1464
              |||   ||||||   |||      ||||||  ||||||                 :::|||
Db        405 Ile---AlaGlyAlaProGlyPheProGlyAlaArgGly-----------ProSerGly 419

Qy       1463 TCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACAC 1404
                   ||||||   |||  ||||||     |||   |||        |||
Db        420 Pro-GlnGlyProGlyGlyProProGlyProLysGlyAsnSerGlyGluProGly---Al 438

Qy       1403 TGTGGGACAGGCATGTGGCACCGGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGA 1344
                  |||            |||||| |||   |||:::
Db        438 aProGlySerLysGlyAspThrGly-AlaLysGlyGluProGlyPro------------ 453
```

```
Qy    1343 CTGCTCGAGTGCCGAATCGCTGCACCAGCCGGTCCATGACCAGAG------AGAAGACCA 1290
                       ||| :::     |||   ||||||      |||         :::|||
Db     454 --------Val-GlyValGlnGlyProProGlyProAlaGlyGluGluGlyLysArgGly 470

Qy    1289 GGGAGATGGCGCACTGCAGGAACAGCC---CCAGGCTGCCCATCCGAACGC--------- 1242
              |||               |||      ||||||   |||    ||||||
Db     471 AlaArgGlyGluProGlyProThrGlyLeuProGlyProProGlyGluArgGlyGlyPro 490

Qy    1241 ------------------CTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTC 1200
                                         |||:::|||||||:::   ||||||
Db     491 GlySerArgGlyPheProGlyAlaAspGlyValAlaGlyProLysGlyProAlaGlyGlu 510

Qy    1199 TGGGCACGCCCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCA 1140
                    ||||||   |||            :::|||
Db     511 ArgGlySerProGlyProAlaGlyProLysGlySerPro--------------------- 523

Qy    1139 TGAGTGCCATCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGC 1080

Db     523 ------------------------------------------------------------ 523

Qy    1079 GGCAGCACAGCTGGTGCAGCCGGGGAAGCAGGGCGCCCAGGT-----------TCCGGA 1032
                    ||||||      ||||||||||||   ||||||            :::|||
Db     524 ---GlyGluAlaGlyArgProGlyGluAlaGlyLeuProGlyAlaLysGlyLeuThrGly 542

Qy    1031 AAGCCAAGCGGGCCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTT 972
              |||      ||||||   |||               :::   ||||||
Db     543 SerProGlySerProGlyProAspGly-----------LysThrGlyProPro------ 556

Qy     971 CTGCTGGCTCGGTGGGGCCCAGCG------------------CTGCCTCCTCAGCCACCA 930
                          ||||||||                      ||||||
Db     557 --------------GlyProAlaGlyGlnAspGlyArgProGlyProProGlyProPro 571

Qy     929 -----------------------------------------------GCAGTGTGGCTG 918
                                                           |||
Db     572 GlyAlaArgGlyGlnAlaGlyValMetGlyPheProGlyProLysGlyAlaAla------ 589

Qy     917 CTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGCACT------CCTCCTGGG 864
                       |||      |||   ||||||:::||||||         |||||||||
Db     590 -----------GlyGluProGlyLysAlaGlyGluArgGlyValProGlyProProGly 605

Qy     863 TGCCCAGGT-------------------------------------------------AGG 852
                 |||
Db     606 AlaValGlyProAlaGlyLysAspGlyGluAlaGlyAlaGlnGlyProProGlyProAla 625

Qy     851 GGGCCAGGGCACTGGTGTCCCAGTCAATGGCAGGCAGGA--------------------- 813
           ||||||                  |||      ||||||
Db     626 GlyProAlaGlyGluArgGlyGluGlnGlyProAlaGlySerProGlyPheGlnGlyLeu 645

Qy     812 ---GGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGC 756
              |||      |||   |||   :::                  |||      ||||||
Db     646 ProGlyProAlaGlyProProGlyGluAlaGlyLysProGlyGluGlnGlyValProGly 665

Qy     755 GACAGTGGT---CCGGGTCCC------GGAAGAGGTCAGAGAGCAGGGCCTCCAGTGGAG 705
           |||   |||   |||||||||         ||||||:::|||   |||   |||   |||
Db     666 AspLeuGlyAlaProGlyProSerGlyAlaArgGlyGluArg---GlyPheProGlyGlu 684

Qy     704 TGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGCCCAGGATGAGCAGTGCCA--- 648
                       |||         ||||||   ||||||               |||
Db     685 ArgGlyValGlnGlyProProGlyProAlaGlyProArgGlyAlaAsnGlyAlaProGly 704

Qy     647 --------------------------GCTCCAGGGGCCTGGGATCCGGGCACAGCAGCC 615
                                      ||||||||||||      ||||||:::
Db     705 AsnAspGlyAlaLysGlyAspAlaGlyAlaProGlyAla------ProGlySerGlnGly 722

Qy     614 CTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCAAGGACAGTG 555
                       :::   :::      |||||||||:::||||||   |||
Db     723 AlaProGlyLeuGlnGlyMetProGlyGluArgGlyAlaAlaGlyLeuProGly------ 740

Qy     554 CCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTA 495
           |||:::         |||                  |||   |||         |||
Db     741 ProLysGlyAspArgGlyAspAlaGlyProLysGlyAlaAspGly---------SerPro 757

Qy     494 GGAGCG---GGACACAGA-----------------------CCAGGCCCAGCACTGGAC 462
           |||      |||   |||                        |||||||||||
Db     758 GlyLysAspGlyValArgGlyLeuThrGlyProIleGlyProProGlyProAla------ 775

Qy     461 CAATGCCCA--------------------------------------------------- 453
                 |||
Db     776 GlyAlaProGlyAspLysGlyGluSerGlyProSerGlyProAlaGlyProThrGlyAla 795

Qy     452 ------GCACCATGGTCATGAACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCA 399
                 ||||||               |||   ||||||   ||||||      |||
Db     796 ArgGlyAlaProGlyAspArgGlyGluProGlyProProGlyProAlaGlyPheAlaGly 815

Qy     398 CATAGGTGATGCCTGCGGCCAAACACACCTCCAGGC------------CAAAGGTTAGCA 351
                                      |||||||||
Db     816 --------------------------ProProGlyAlaAspGlyGlnProGlyAlaLys 826

Qy     350 GGTTGACCAGCAAGAGCTGGGCTTTCCGGTGCCGCAGCAGGCGGCTCACCCACAGCCTCT 291
           |||   |||      ||||||      |||   |||      |||   |||      :::
Db     827 GlyGluProGlyAspAlaGlyAlaLysGlyAspAla------GlyProProGlyProAla 844

Qy     290 GGACCATAGTGGGCCA------------GGCGGGTAGGGCTCAGGGGGCCGTTCAGGCAC 243
           |||||| :::||||||           |||   |||:::   |||   |||
Db     845 GlyPro-AlaGlyProProGlyProIleGlyAsnValGlyAlaProGlyAlaLysGlyAl 864

Qy     242 TCCAGAACTGCTTCGTCTCGGCTCTGCT------CCAGAAGCTGCGGCCTCTCCTCCTTG 189
                              ||||||||       |||   |||         |||
Db     864 a---------------ArgGlySerAlaGlyProProGlyAlaThrGlyPheProGlyAl 879

Qy     188 CTGCCGCCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGG 129
                  ::::::            |||      :::         |||         ||
Db     879 aAlaGlyArgValGlyProProGlyProSerGlyAsnAlaGlyProProGlyProProGl 899

Qy     128 TCCAGCTTCTCAGCCCATGCTCAACACCTGCTGCTGTGGG------GCACCTCAGTGGGG 75
           |||||||   :::                        |||         |||:::  ||:
Db     899 yProAlaGlyLysGlu--------------------GlyGlyLysGlyProArg-GlyG 912

Qy      74 ACACGTCTCATCACTCAGATCCTGGCCGA-----------------------------G 45
           ::|||               ||||||                                |
Db     912 luThrGlyProAlaGlyArgProGlyGluValGlyProProGlyProProGlyProAlaG 932

Qy      44 GCGCGCGGCTGTCACCCGGAGCC 22
           ||   :::   |||||||||||
Db     932 lyGluLysGlySerProGlyAla 939
```

```
RESULT 41
US-10-096-986-103
; Sequence 103, Application US/10096986
; Publication No. US20030083464A1
;   GENERAL INFORMATION:
;        APPLICANT: Ferrari, Franco A.
;                   Richardson, Charles
;                   Chambers, James
;                   Causey, Stuart
;                   Pollock, Thomas J.
;                   Cappello, Joseph
;                   Crissman, John W.
;        TITLE OF INVENTION: No. US20030083464Alel Peptides Comprising Repetitive
;                      Units of Amino Acids and DNA Sequences Encoding the Same
;        NUMBER OF SEQUENCES: 117
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Flehr Hohbach Test Albritton & Herbert LLP
;             STREET: Four Embarcadero Center, Suite 3400
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: US
;             ZIP: 94111
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/096,986
```

```
;            FILING DATE: 12-Mar-2002
;            CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/444,791
;            FILING DATE: 22-No. US20030083464A1-1999
;            APPLICATION NUMBER: US 08/482,085
;            FILING DATE: 07-JUN-1995
;            APPLICATION NUMBER: US 08/175,155
;            FILING DATE: 29-DEC-1993
;            APPLICATION NUMBER: US 08/053,049
;            FILING DATE: 22-APR-1993
;            APPLICATION NUMBER: US 07/114,618
;            FILING DATE: 29-OCT-1987
;            APPLICATION NUMBER: US 06/927,258
;            FILING DATE: 04-NOV-1986
;         ATTORNEY/AGENT INFORMATION:
;            NAME: Trecartin, Richard F.
;            REGISTRATION NUMBER: 31,801
;            REFERENCE/DOCKET NUMBER: A-55186-11/RFT/BTC
;         TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 415-781-1989
;            TELEFAX: 415-398-3249
;      INFORMATION FOR SEQ ID NO: 103:
;         SEQUENCE CHARACTERISTICS:
;            LENGTH: 837 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;         MOLECULE TYPE: peptide
;         SEQUENCE DESCRIPTION: SEQ ID NO: 103:
US-10-096-986-103

Alignment Scores:
Pred. No.:              1.83e-10      Length:         837
Score:                  323.00        Matches:        280
Percent Similarity:     30.52%        Conservative:   45
Best Local Similarity:  26.29%        Mismatches:     354
Query Match:            5.03%         Indels:         386
DB:                     9             Gaps:           59

US-09-759-143-110 (1-3410) x US-10-096-986-103 (1-837)

Qy     84 GGTGCCCCACAGCAGCAGGTGTTG------AGCATGGGCTGAGAAGCTGGACCGGCACCA 137
Db     37 GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro 56

Qy    138 AAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGCGGCAGCAAGGAGGA 197
Db     57 GlnGly-----------LeuProGly------------------------------ 61

Qy    198 GAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAGTGCCTGAACGGCCC 257
Db     62 ---------SerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro 78

Qy    258 CCTGAGCCCTACCCGCCTGGCCCACTATGGT-----------CCAGAGGCTGTGGGTGA 305
Db     79 GlyAlaProGlyThrProGlyPro-GlnGlyLeuProGlySerProGlyAlaProGlyTh 98

Qy    306 GCCG---CCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTG 362
Db     98 rProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly-------- 115

Qy    363 GCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT-- 420
Db    116 -ProGlnGlyLeuProGly------------------SerProGlyAlaProGlyThrPr 129

Qy    421 -GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCT-- 477
Db    129 oGlyProGlnGlyLeuProGlySerProGlyAlaPro------GlyThrProGlyProGl 147

Qy    478 -GGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCC 536
Db    147 nGlyLeuProGlySerPro--------------------------------------- 153

Qy    537 GCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAA 596
Db    154 --------------------------GlyAlaProGlyThrPro------GlyProGl 162

Qy    597 GGGCCGGCTGGCTAGCAGGGCTGCT------GTGCCCGGATCCCAGGCC---CCTGGAGC 647
Db    162 nGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySe 182

Qy    648 TGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTC 707
Db    182 rProGlyAlaProGlyThrProGlyProGlnGlyLeu------ProGlySerProGlyAl 200

Qy    708 CACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGA---CCCGGA---CCACTGTCGCCAGG 761
Db    200 aProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGl 220

Qy    762 CCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTG 821
Db    220 y-------------------------ProGlnGlyLeuProGlySerProGlyAla-- 230

Qy    822 CCATTGACTGGGACACCAGTGCCCTGGCCCC------CTACCTGGGCACCCAGGA----- 870
Db    231 ------------ProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPr 246

Qy    871 ------------------GGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCT 911
Db    246 oGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPr 266

Qy    912 GCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAG 971
Db    266 oGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGl 286

Qy    972 AAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTT 1031
Db    286 ySerProGlyAlaPro--------------------GlyThrProGlyProGlnGlyLe 299

Qy   1032 TCCG---GAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCC 1088
Db    299 uProGlySerProGlyAlaProGlyThrPro---GlyProGlnGlyLeuProGlySerPr 318

Qy   1089 GCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCA 1148
Db    318 o---GlyAlaProGlyThrProGly---------------------------------- 325

Qy   1149 CGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGC 1208
Db    326 -------ProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro----------- 339

Qy   1209 CGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCT-- 1266
Db    340 ----------GlyProGlnGlyLeu---------------ProGlySerProGlyAlaPr 351

Qy   1267 ----GTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGC 1322
Db    351 oGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly-- 370

Qy   1323 AGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCG 1382
Db    371 ------------ProGlnGlyLeuProGlySerProGlyAla---Pro--GlyThrPro 384

Qy   1383 GTGCCACATGCCTGTCCC--------------------ACAGTGTGGCCGTGGTGACAG 1421
Db    385 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro------Gln 402

Qy   1422 CTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGG 1481
Db    403 GlyLeuProGlySerProGlyAlaProGly-ThrProGlyPro-----------GlnGl 418

Qy   1482 CCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAA------ATACCGAGGGACA 1535
Db    418 yLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuPro-GlySerP 438
```

```
Qy     1536 CTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAG---- 1591
              |||        :::                            |||||||||:::
Db      438 roGlyAlaProGlyThr-------------------------ProGlyProGlnGlyL 449

Qy     1592 -------------------CCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAG 1631
                               ||||||  |||   |||   |||    |||:::   |
Db      449 euProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProG 469

Qy     1632 GCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGT------GATG 1685
            ||:::            |||   |||      |||    |||:::
Db      469 lyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrP 489

Qy     1686 TCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCA 1745
               :::   :::   |||  ||||||       |||  |||      ||||||
Db      489 roGlyProGlnGlyLeuProGly-SerProGlyAlaPro-GlyThrProGlyProGlnGl 508

Qy     1746 TCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCC 1805
             ||||||||                     |||   ||||||   |||    ||
Db      508 yLeuProGlySer--------------------ProGlyAlaProGlyThrProGlyPr 521

Qy     1806 TGTTTATGGGCTC-----------CATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATA 1853
            |      ||||||                |||      |||         |||
Db      521 o---GlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuPr 540

Qy     1854 TGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTG 1913
              |||            ||||||:::
Db      540 oGlySerProGlyAlaProGlyThrPro------------------------------- 549

Qy     1914 ACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGG 1973
                                                            |||   ::::||
Db      550 -----------------------------------------------GlyProGlnGl 553

Qy     1974 CCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGC 2033
             ||||||            ||| |||           |||   |||||||||||    |
Db      553 yLeuProGlySerProGlyAla-ProGlyThrProGlyProGlnGlyLeuProGlySerP 573

Qy     2034 CGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAG 2093
             ||                                                      |
Db      573 ro-------------------------------------------------------G 574

Qy     2094 GTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGG 2153
             |||||      |||   |||    |||:::|||      ||||||   :::    |||
Db      574 lyAla-ProGlyThrProGlyProGlnGlyLeuPro---GlySerProGlyAlaProGly 592

Qy     2154 CTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAA 2213
               |||       :::           |||              |||      |||:::
Db      593 ThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGln 612

Qy     2214 GGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT----------------------- 2249
             ||||||                |||   |||
Db      613 GlyLeuProGly--------------SerProGlyAlaProGlyThrProGlyProGln 627

Qy     2250 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCT---AGAGAA 2306
             |||||||||                    |||   .:::    ||||||       :::
Db      628 GlyLeuPro---------------GlySerProGlyAlaProGlyThrProGlyProGln 642

Qy     2307 GGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACC 2366
             |||     |||              |||||||||   |||      ::: |||   ||
Db      643 GlyLeuProGly--------------SerProGlyAlaProGly---Thr-ProGlyPr 656

Qy     2367 TGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATG 2426
            |
Db      656 oGlnGly--------------------------------------------------- 658

Qy     2427 GGATTTGAACATATGACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACC 2486
                                                  |||    |||:::        ||
Db      659 -----------------------------------LeuProGlySerProGlyAlaPr 666

Qy     2487 AGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATC 2546
             ||||      ||||||
Db      666 oGlyThrProGlyProGlnGly------------------------------------ 673

Qy     2547 AGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTA 2606

                                    ||||||
Db      674 -----------------------LeuPro----------------------------- 675

Qy     2607 ACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTA 2666

Db      675 ------------------------------------------------------------ 675

Qy     2667 ATATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGG 2726
                    |||     |||    |||          ||||||
Db      676 -------GlySerProGlyAlaProGlyThrProGlyPro------------------- 686

Qy     2727 GCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAG 2786
                                          ||||||              |||
Db      687 --------------------------GlnGlyLeuProGlySerProGlyAlaProGl 697

Qy     2787 GACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGG 2846
              |||              :::|||                        |||
Db      697 yThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro-------- 714

Qy     2847 GTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTC 2906
                                  |||     ||||||   |||   |||     |||
Db      715 ---------------------GlyProGlnGlyLeuProGlySerProGlyAlaProGl 727

Qy     2907 TTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA---CTCCCCTCTACTCTCTCTAGGAC 2963
              :::   ||||||        |||      |||         |||     |||
Db      727 yThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGl 747

Qy     2964 TGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAAC 3023
              |||           ||||||            |||       |||   |||
Db      747 nGly-----------LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGl 763

Qy     3024 TTTCCCCACCAGCTCCACAACCCTGTTTGGAGCT------ACTGCAGGACCAGAA 3072
                |||   |||               ||||||      |||    ||||||:::
Db      763 yLeuProGlySerPro------------GlyAlaProGlyThrProGlyProGln 777
```

RESULT 42
US-09-919-497-56
; Sequence 56, Application US/09919497
; Patent No. US20020106662A1
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 56
; LENGTH: 1806
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (758)..(758)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: UNSURE
; LOCATION: (809)..(809)
; OTHER INFORMATION: Xaa = any amino acid
US-09-919-497-56

```
Alignment Scores:
Pred. No.:                4.38e-10      Length:         1806
Score:                    316.00        Matches:        315
Percent Similarity:       31.07%        Conservative:   61
Best Local Similarity:    26.03%        Mismatches:     413
Query Match:              5.08%         Indels:         423
DB:                       10            Gaps:           71

US-09-759-143-110 (1-3410) x US-09-919-497-56 (1-1806)

Qy     3101 ACAAAGGCTTGGGAAACCGCACTTTGTGCTTCTGGT---------------------- 3066
```

```
                 |||           ||||||::::::        |||
Db        409 ThrAspIleThrGluThrSerIleAsnGlyHisGlyAlaTyrGlyGluLysGlyGlnLys 428

Qy       3065 ------CCTGCAGTAGCT---CCAAACAGGGTTGTGGAGCTGGTGGGGAAAGTTGGGGGT 3015
                    |||||||||    |||     ::::||||||            |||
Db        429 GlyGluProAlaValValGluProGlyMetLeuValGlu---------------GlyPro 443

Qy       3014 AGGGGAAAGTTGGGGGTAGGGGAAATTTTGGGCAGTGCCTTCATCAGCCCAGTCCTAGAG 2955
                 |||      |||        |||:::|||               |||    |||:::
Db        444 ProGlyProAlaGlyProAlaGlyIleMetGly-----------ProProGlyLeuGln 459

Qy       2954 AGAGTAGAGGGGAGTGGAAGTGGGGGGAACCAGGCTGGGCCAAGAGAAGAGGGGTGGTTA 2895
                                  |||      |||      |||   :::   |||
Db        460 ---------------GlyProThrGlyProProGlyAspProGlyAspArgGlyProPro 474

Qy       2894 GGGAAGCCGTTGAGACCTGAAGCC---------------CCACCCTCTACCTTCCTTCAA 2850
              |||:::|||      |||   |||              ||||||   |||    |||
Db        475 GlyArgProGlyLeuProGlyAlaAspGlyLeuProGlyProProGlyThrMetLeuMet 494

Qy       2849 CACCCTAACCTTGGGTAACAGCATTTGGAATTATCATTTGGGATGAGTAGAATTTCCAAG 2790
                 |||                                   :::|||
Db        495 LeuPro-----------------------PheArgTyrGlyGlyAsp------------ 502

Qy       2789 GTCCTGGGTTAGGCATTTTGGGGGGCCAGA---------------CCCCAGGAGAAGAAG 2745
                                       |||::::::              ||||||   :::
Db        503 ---------------------GlySerLysGlyProThrIleSerAlaGlnGluAlaGln 515

Qy       2744 ATTCTGGCAATGATCAGCCCAATGACCAGCTATCTCAGGGGACCTGAT------------ 2697
                    |||::::::                 ||||||||||||
Db        516 AlaGlnAlaIleLeuGlnGlnAlaArgIleAlaLeuArgGlyProProGlyProMetGly 535

Qy       2696 ---TGTTGGGGATCCCCCACCCTACCCAAATATTAGACACCAACACAGAAAAGCTAGCAA 2640
                    |||   |||   |||                        |||       ::
Db        536 LeuThr-GlyArgProGlyProValGlyGlyProGlySerSerGlyAlaLysGlyGluSe 555

Qy       2639 TGGATTCCCTTCTACTTTGTTAAATAAATAAGTTAAATATTTAAATGCCTGTGTCTCTGT 2580
              :|||   |||                                             ||
Db        555 rGlyAspProGlyPro--------------------------GlnGlyProArgGlyVa 566

Qy       2579 GATGGCAACAGAAGGACCAACAGGCCACATCCTGATAAAAGG------------------ 2538
              |            ||||||||||||     |||    ||||||
Db        566 lGlnGlyProProGlyProThrGly--LysProGlyLysArgGlyArgProGlyAlaAsp 585

Qy       2537 ------------------TAAGAGGGGGGTGGATCAGCAAAAAGACAGTGCTGTGGGCTG 2496
                                   |||     |||           |||          ||||||
Db        586 GlyGlyArgGlyMetProGlyGluProGlyAlaLysGlyAspArgGlyPheAspGlyLeu 605

Qy       2495 AGGGGA---CCTGGTTCTTGTGTGTTGCCCCTCAGGACTCTTCCCCTACAAATAAGTCAT 2439
                 |||    ||||||
Db        606 ProGlyLeuProGlyAsp------------------------------------------ 611

Qy       2438 ATGTTCAAATCCCATGGAGGAGTGTTTCATCCTAGAAACTCCCATGCA------------ 2391
                    |||    |||    |||          |||:::
Db        612 ------LysGlyHisArgGlyGluArgGlyProGlnGlyProProGlyProProGlyAsp 629

Qy       2390 ------------------------------------AGAGCTACATTAAACGAAGCTGCA 2367
                                                     |||          ||||||
Db        630 AspGlyMetArgGlyGluAspGlyGluIleGlyProArgGlyLeuProGlyGluAlaGly 649

Qy       2366 GGTTAAGGGGCTTAGAGAT------GGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAA 2313
                :::||||||               |||    ||||||              ||||||
Db        650 -ProArgGlyLeuLeuGlyProArgGlyThrProGly---------------AlaProGl 664

Qy       2312 AAACCCTTCTCTA---GGTGTGTCTCAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGG 2256
                  |||   :::   ||||||             |||                   ||
Db        664 yGlnProGlyMetAlaGlyValAspGlyProProGly---------------ProLysGl 679

Qy       2255 TAAT------------------------CCACCTGCAGAGTCCCCGCATTCCAGTGCATG 2220
              ||||                         ||||||    :::       :::
Db        679 yAsnMetGlyProGlnGlyGluProGlyProProGlyGlnGlnGlyAsnPro-------- 696

Qy       2219 GAGCCCTTCTGGCCTCCCT---GTATAAGTCCAGACTGAAACCCCCTTGGAAGGCCTCCA 2163
                  |||   |||||||||         :::        :::    ||||||    :::

Db        697 -GlyProGlnGlyLeuProGlyPro-GlnGlyProIleGlyProProGlyGluLysGlyP 716

Qy       2162 GTCAG---------------GCAGCCCTA------------------------------ 2149
               |||               |||  |||
Db        716 roGlnGlyLysProGlyLeuAlaGlyLeuProGlyAlaAspGlyProProGlyHisProG 736

Qy       2148 -----------GAGACTGGGGAGAGAGGAGAG---GGACGCCCCAGCCCCCAGCTGTGCA 2103
                          ::::::||||||::::|||     |||    |||    ||||||
Db        736 lyLysGluGlyGlnSerGlyGluLysGlyAlaLeuGlyProProGlyProGlnGlyProI 756

Qy       2102 GCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGCAACAG 2043
                       |||          |||                 :::
Db        756 leGly***ProGlyProArgGlyValLysGlyAlaAspGlyValArgGlyLeuLysGlyS 776

Qy       2042 AAACTGGC------GGCCAGCCCGGCAGCCCC---------------ATGGGGCTAACAG 2004
                 |||       |||:::   |||   |||                |||||||||  |
Db        776 erLysGlyGluLysGlyGluAspGlyPheProGlyPheLysGlyAspMetGlyLeuLysG 796

Qy       2003 GAGC-------------GGGGAGCTGGGACCC---------------------------- 1985
               ||               |||::::::||||||
Db        796 lyAspArgGlyGluValGlyGlnIleGlyProArgGly***AspGlyProGluGlyProL 816

Qy       1984 --AGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGCC 1927
                    |||||||||              |||                   ||||
Db        816 ysGlyArgAlaGlyProThrGlyAspProGlyProSer------------------GlyG 830

Qy       1926 AAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAGAC-------- 1875
              ||                                       |||::::::
Db        830 lnAla-----------------------------------GlyGluLysGlyLysLeuG 838

Qy       1874 -----CCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTGGACAAT 1820
                  |||||||||      |||    ||||||
Db        838 lyValProGlyLeuProGlyTyrProGlyArgGln------------------------ 849

Qy       1819 GGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCACTATCCAGGATGG- 1761
                                   ||||||   |||:::    |||       :::||||||:::
Db        850 -------------------GlyProLysGlySerThrGly------PheProGlyPheP 861

Qy       1760 -------------------------CGAGGTCCAGGCAGATGCCCCGGCCCGGAA---- 1731
                                        ||||||            |||||||||
Db        861 roGlyAlaAsnGlyGluLysGlyAlaArgGlyValAlaGlyLysProGlyProArgGlyG 881

Qy       1730 --------CCACCCTGGCCTCGGTGGGCTCACCCACCACCACACGTACGGAGACATCACA 1679
                 |||       ||||||   ||||||               |||     |||   ::
Db        881 lnArgGlyProThrGlyProArg--GlySer---ArgGlyAlaArgGlyProThrGlyLy 899

Qy       1678 GGCAGAGGCCCCGCAGAGCGCGGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCAC 1619
              :        |||:::              ||||||          |||   ||||||
Db        899 s---ProGlyProLysGlyThrSerGlyGlyAspGlyProProGlyProProGlyGluAr 918

Qy       1618 GTGTCCATTAGGGAAGGGAGCTCCAGGCTTAGGGCCTGGCAGGAAGCTGGTCATCAGGCT 1559
                 |||   |||        |||         ||||||   |||
Db        918 gGlyProGlnGlyProGlnGlyProValGlyPheProGlyProLys-------------- 933

Qy       1558 GTCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTC 1499
                               ||||||      |||  |||          |||     |||     |
Db        934 ------------GlyProProGlyProPro-Gly---ArgMetGlyCysProGlyHisP 948

Qy       1498 CCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCC 1439
              |||||  ||||||                |||  ||||||        |||   |||
Db        948 roGlyGlnArgGly-----------GluThrGly-PheGlnGlyLysThrGlyProPro 963

Qy       1438 GGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACCGGC 1379
               |||    ||||||                       ||||||            ||||||
Db        964 GlyProGlyGly-------------------ValValGlyProGlnGlyProThrGly 976

Qy       1378 AGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCTGCAC 1319
                 ||||||                                        |||  |||
Db        977 Glu-ThrGly---------------------------------Pro-IleGlyGluA 984

Qy       1318 CAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACAGCC---- 1263
                       |||    |||       ||||||              |||   |||
Db        984 rgGlyTyrProGlyProProGlyProProGlyGluGlnGlyLeuProGlyAlaAlaGlyL 1004
```

```
Qy        1262 --CCAGGCTGCCCATCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTC 1205
                  |||            :::            ::::||||||          ||||||
Db        1004 ysGluGlyAlaLysGlyAspProGlyProGlnGlyIleSerGlyLysAspGlyProAlaG 1024

Qy        1204 AGCTCTGGGCACGCCCTGGTACA-----------------------GCCCCT------- 1176
                  |||         ||||||                          ||||||
Db        1024 lyLeuArgGlyPheProGlyGluArgGlyLeuProGlyAlaGlnGlyAlaProGlyLeuL 1044

Qy        1175 ----------CGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCA 1127
                         |||:::   |||                               |||
Db        1044 ysGlyGlyGluGlyProGlnGlyPro---------------------------ProGlyP 1055

Qy        1126 GCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACAGCTG 1067
                     :::         |||     ||||||                       ||||||
Db        1055 roValGlySerProGlyGluArgGlySerAlaGly------------------ThrAlaG 1069

Qy        1066 GTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGGCCCGGCATGGACA 1007
                ||       |||      |||||||||||          |||    ||||||       :::
Db        1069 lyProIleGlyLeuArgGlyArgProGlyProGlnGlyProProGlyProAlaGlyGluL 1089

Qy        1006 GCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCGC 947
                                 :::||||||                          ||||||||
Db        1089 ysGlyAlaProGlyGluLysGlyProGln---------------------GlyProAla- 1101

Qy         946 TGCCTCCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAG 887
                                                          ||||||   |||      |
Db        1102 ------------------------------------------GlyArgAspGlyValGlnG 1108

Qy         886 GCCAAAGAGGCACTCCTCCTGGGTGCCCAGGT---AGGGGGCCAGGGCACTGGTGTCCCA 830
                ||           |||   |||   ||||||    ||||||
Db        1108 ly-----------ProValGlyLeuProGlyProAlaGlyPro---------------- 1118

Qy         829 GTCAATGGCAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCATAGAC 770
                              |||||||||||         |||             :::
Db        1119 --------------AlaGlySerProGlyGluAspGlyAsp---------LysGlyGluI 1131

Qy         769 AGAGTAGGCCTGGCGACAGTGGTCCGGGTCCCGGAAGAGGTCAGAGAGCAGGGCCTCCAG 710
                        ||||||       |||          |||:::|||:::      |||||||||
Db        1131 leGlyGluProGlyGlnLysGlySerLysGlyGlyLysGlyGluAsn---GlyProPro- 1149

Qy         709 TGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGCCCAGGATGAGCAGTGC 650
                          ||||||    :::    |||          ||||||
Db        1150 --------GlyProProGlyLeuGlnGlyProValGly---AlaProGly---------I 1163

Qy         649 CAGCTCCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGAT 590
                  |||    |||      :::||||||
Db        1163 leAlaGlyGlyAspGlyGluProGly------------------------------- 1171

Qy         589 GAGAAAGAGGCTCAGCAGGATGCCCAAGGACAGTGCCCAGATGAAGGGCCGGCGGCGGCC 530
                    |||||| ||||||            ||||||      |||||||||   |||
Db        1172 --ProArgGly-GlnGlnGlyMetPheGlyGlnLysGlyAspGluGly---AlaArgGly 1189

Qy         529 ATAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAG 470
                        |||  |||             ::::|||            |||||||||
Db        1190 PheProGlyProPro--Gly---------ProIleGlyLeuGlnGlyLeuProGlyProP 1206

Qy         469 CACTGGACCAATGCCCA-----------GCACCATGGTCATGAACTTCTCCTCTACCCC 422
                       ::::::                 ||||||         |||    ||||
Db        1206 roGlyGluLysGlyGluAsnGlyAspValGlyProTrpGly------ProProGlyProP 1224

Qy         421 CACTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCCAAACACACCTCCAGGCC 362
                ||    |||                               |||      |||||||||
Db        1224 roGlyProArgGlyProGlnGlyProAsnGlyAlaAspGlyProGlnGlyProProGly- 1243

Qy         361 AAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGTGCCGCAGCAGGCGGCTCAC 302
                                     :::         :::   |||             |||    |
Db        1244 --------------------SerValGlySerValGlyGlyValGlyGluLysGlyGluP 1257

Qy         301 CCACAGCCTCTGGACCATAGTGGG---CCAGGCGGGTAGGGCTCAGGGGGCCGTTCAGGC 245
                ||     :::|||  :::     |||   ||||||         |||     ||||||        |||
Db        1257 roGlyGluAlaGly-AsnProGlyProProGlyGluAlaGlyValGlyGlyProLysGly 1276

Qy         244 ACTCCAGAACTGCTTCGTCTCGGCTCTGCT---CCAGAAGCTGCGGCCTCTCCTCCTTGC 188
                      |||                  |||   |||    |||        ||||||   ||||||
Db        1277 ------GluArgGlyGluLysGlyGluAlaGlyProProGlyAlaAlaGlyProProGly 1294

Qy         187 TGCCGC---------------CAACTGCCTAGGAATCAGCCAGGCGCCCATTTCT----- 148
                :::                              |||:::     ||||||  |||:::
Db        1295 AlaLysGlyProProGlyAspAspGlyProLysGlyAsnProGly-ProValGlyPhePr 1314

Qy         147 -GCCAGCCCTTTGGTGCCGGTCCAGCTTCTCAGCCCATGCTCAACACCTGCTGCTGTGGG 89
                      |||        |||    :::|||                        ||||||
Db        1314 oGlyAspProGlyProProGlyGluLeu---------------GlyProAlaGlyGlnAs 1329

Qy          88 GCACCTCAGTGGGGACACGTCTCATCACTCAGATCCTGGCCGA----------------- 46
                     :::    ||||||                    |||||||||:::
Db        1329 pGlyValGlyGlyAspLysGlyGluAspGlyAspProGlyGlnProGlyProProGlyPr 1349

Qy          45 ------------------------------GGCGCGCGGCTGTCACCCGGAGCCAGC 19
                                                |||      |||         |||||||||:::
Db        1349 oSerGlyGluAlaGlyProProGlyProProGlyLysArgGlyProProGlyAlaAla 1368
```

RESULT 43
US-10-012-896-547
; Sequence 547, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Wantanabe, Yoshihiro
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012,896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 547
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-896-547

Alignment Scores:			
Pred. No.:	8.01e-10	Length:	58
Score:	312.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.86%	Indels:	0
DB:	9	Gaps:	0

US-09-759-143-110 (1-3410) x US-10-012-896-547 (1-58)

```
Qy         932 GTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCC 991
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ValAlaGluGluAlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSer 20

Qy        992 TTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTG 1051
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuSerProHisCysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeu 40

Qy       1052 CTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTC 1105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 LeuProArgLeuHisGlnLeuCysCysArgMetProArgThrLeuArgArgLeu 58
```

RESULT 44
US-09-895-793-547
; Sequence 547, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 547
;   LENGTH: 58
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-895-793-547

```
Alignment Scores:
Pred. No.:               8.01e-10       Length:        58
Score:                   312.00         Matches:       58
Percent Similarity:      100.00%        Conservative:  0
Best Local Similarity:   100.00%        Mismatches:    0
Query Match:             4.86%          Indels:        0
DB:                      9              Gaps:          0
```

US-09-759-143-110 (1-3410) x US-09-895-793-547 (1-58)

```
Qy        932 GTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCC 991
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ValAlaGluGluAlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSer 20

Qy        992 TTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTG 1051
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuSerProHisCysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeu 40

Qy       1052 CTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTC 1105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 LeuProArgLeuHisGlnLeuCysCysArgMetProArgThrLeuArgArgLeu 58
```

RESULT 45
US-09-895-814-547
; Sequence 547, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 547
;   LENGTH: 58
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-895-814-547

```
Alignment Scores:
Pred. No.:               8.01e-10       Length:        58
Score:                   312.00         Matches:       58
Percent Similarity:      100.00%        Conservative:  0
Best Local Similarity:   100.00%        Mismatches:    0
Query Match:             4.86%          Indels:        0
DB:                      9              Gaps:          0
```

US-09-759-143-110 (1-3410) x US-09-895-814-547 (1-58)

```
Qy        932 GTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCC 991
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ValAlaGluGluAlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSer 20

Qy        992 TTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTG 1051
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuSerProHisCysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeu 40

Qy       1052 CTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTC 1105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 LeuProArgLeuHisGlnLeuCysCysArgMetProArgThrLeuArgArgLeu 58
```

Search completed: June 9, 2003, 22:57:11
Job time : 249 secs

```
OM nucleic - nucleic search, using sw model

Run on:         June  9, 2003, 18:34:24 ; Search time 442 Seconds
                                          (without alignments)
                                          10797.336 Million cell updates/sec

Title:          US-09-759-143-110
Perfect score:  3410
Sequence:       1 gggaaccagcctgcacgcgc..........aaaaaaataaaaaaaaaaaa 3410

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       870385 seqs, 699768693 residues

Total number of hits satisfying chosen parameters:      1740770

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_NA:*
                 1:  /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
                 2:  /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
                 3:  /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
                 4:  /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
                 5:  /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
                 6:  /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
                 7:  /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
                 8:  /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
                 9:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
                 10:  /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
                 11:  /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
                 12:  /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
                 13:  /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
                 14:  /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3409.6	100.0	3410	9	US-09-232-880-110	Sequence 110, App
2	3409.6	100.0	3410	9	US-10-012-896-110	Sequence 110, App
3	3409.6	100.0	3410	9	US-09-895-793-110	Sequence 110, App
4	3409.6	100.0	3410	9	US-09-895-814-110	Sequence 110, App
5	3409.6	100.0	3410	9	US-10-010-940-110	Sequence 110, App
6	3409.6	100.0	3410	10	US-09-745-288-100	Sequence 100, App
7	3409.6	100.0	3410	10	US-09-759-143-110	Sequence 110, App
8	3409.6	100.0	3410	10	US-09-780-669-110	Sequence 110, App
9	3409.6	100.0	3410	10	US-09-030-606-110	Sequence 110, App
10	3409.6	100.0	3410	10	US-09-822-827-110	Sequence 110, App
11	3409.6	100.0	3410	10	US-09-115-453-110	Sequence 110, App
12	3292.4	96.6	3320	10	US-09-838-785-1	Sequence 1, Appli
13	2585.4	75.8	4034	9	US-10-012-896-704	Sequence 704, App
14	2585.4	75.8	4034	9	US-09-895-793-704	Sequence 704, App
15	2585.4	75.8	4034	9	US-09-895-814-704	Sequence 704, App
16	2585.4	75.8	4034	10	US-09-759-143-704	Sequence 704, App
17	2585.4	75.8	4034	10	US-09-780-669-704	Sequence 704, App
18	2585.4	75.8	4034	10	US-09-822-827-704	Sequence 704, App
19	2196.4	64.4	2904	9	US-10-012-896-703	Sequence 703, App
20	2196.4	64.4	2904	9	US-09-895-793-703	Sequence 703, App
21	2196.4	64.4	2904	9	US-09-895-814-703	Sequence 703, App
22	2196.4	64.4	2904	10	US-09-759-143-703	Sequence 703, App
23	2196.4	64.4	2904	10	US-09-780-669-703	Sequence 703, App
24	2196.4	64.4	2904	10	US-09-822-827-703	Sequence 703, App
25	2142.8	62.8	4894	9	US-10-012-896-702	Sequence 702, App
26	2142.8	62.8	4894	9	US-09-895-793-702	Sequence 702, App
27	2142.8	62.8	4894	9	US-09-895-814-702	Sequence 702, App
28	2142.8	62.8	4894	10	US-09-759-143-702	Sequence 702, App
29	2142.8	62.8	4894	10	US-09-780-669-702	Sequence 702, App
30	2142.8	62.8	4894	10	US-09-822-827-702	Sequence 702, App
31	1815.8	53.2	6976	9	US-10-012-896-705	Sequence 705, App
32	1815.8	53.2	6976	9	US-09-895-793-705	Sequence 705, App
33	1815.8	53.2	6976	9	US-09-895-814-705	Sequence 705, App
34	1815.8	53.2	6976	10	US-09-759-143-705	Sequence 705, App
35	1815.8	53.2	6976	10	US-09-780-669-705	Sequence 705, App
36	1815.8	53.2	6976	10	US-09-822-827-705	Sequence 705, App
37	794.6	23.3	1203	9	US-10-012-896-851	Sequence 851, App
38	794.6	23.3	1203	9	US-09-895-793-851	Sequence 851, App
39	794.6	23.3	1203	9	US-09-895-814-851	Sequence 851, App
40	794.6	23.3	1203	10	US-09-759-143-851	Sequence 851, App
41	794.6	23.3	1203	10	US-09-780-669-851	Sequence 851, App
42	794.6	23.3	1203	10	US-09-822-827-851	Sequence 851, App
43	673.4	19.7	789	9	US-09-232-880-10	Sequence 10, Appl
44	673.4	19.7	789	9	US-10-012-896-10	Sequence 10, Appl
45	673.4	19.7	789	9	US-09-895-793-10	Sequence 10, Appl

ALIGNMENTS

```
RESULT 1
US-09-232-880-110
; Sequence 110, Application US/09232880
; Publication No. US20020182596A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT: Dillon, Davin C.
;  APPLICANT: Mitcham, Jennifer Lynn
;  TITLE OF INVENTION: COMPOUNDS FOR IMMUNODIAGNOSIS OF
;  TITLE OF INVENTION: PROSTATE CANCER AND METHODS FOR THEIR USE
;  FILE REFERENCE: 210121.428C6
;  CURRENT APPLICATION NUMBER: US/09/232,880
;  CURRENT FILING DATE: 1999-01-15
;  NUMBER OF SEQ ID NOS: 338
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 110
;   LENGTH: 3410
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-232-880-110

  Query Match            100.0%;  Score 3409.6;  DB 9;  Length 3410;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
```

```
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy         301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy         361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy         421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy         481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy         541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy         601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy         661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy         721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy         781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy         841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy         901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy         961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy        1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy        1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy        1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy        1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy        1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy        1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy        1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy        1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy        1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy        1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy        1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy        1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy        1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy        1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy        1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy        1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy        1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy        2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy        2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy        2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy        2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy        2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy        2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy        2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
```

```
Qy       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700

Qy       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy       3361 AAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 AAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
```

```
RESULT 2
US-10-012-896-110
; Sequence 110, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Wantanabe, Yoshihiro
;  APPLICANT:  Meagher, Madeleine Joy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C27
;  CURRENT APPLICATION NUMBER: US/10/012,896
;  CURRENT FILING DATE:  2001-12-10
;  NUMBER OF SEQ ID NOS: 1011
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 110
;   LENGTH: 3410
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-012-896-110

  Query Match             100.0%;  Score 3409.6;  DB 9;  Length 3410;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy     541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy     601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy     661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy     721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy     781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy     841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy     901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy     961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy    1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy    1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy    1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy    1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy    1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy    1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy    1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy    1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy    1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy    1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy    1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy    1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy    1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy    1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy    1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy    1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy    1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy    2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy    2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy    2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy    2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy    2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy    2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy    2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy    2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy    2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy    2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy    2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
```

```
Qy       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
```

```
RESULT 3
US-09-895-793-110
; Sequence 110, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.534C2
;  CURRENT APPLICATION NUMBER: US/09/895,793
;  CURRENT FILING DATE:  2001-06-29
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 110
;   LENGTH: 3410
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-895-793-110

  Query Match             100.0%;  Score 3409.6;  DB 9;  Length 3410;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy       2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy       2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy       2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy       2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700

Qy       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
```

```
Qy       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAA 3410
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAA 3410
```

RESULT 4
US-09-895-814-110
; Sequence 110, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 110
; LENGTH: 3410
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-895-814-110

```
  Query Match             100.0%;  Score 3409.6;  DB 9;  Length 3410;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
```

```
Qy   1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy   1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy   1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy   1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy   1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy   1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy   1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy   1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy   1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy   1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy   1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy   1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy   1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy   1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy   1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy   1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy   2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy   2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy   2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy   2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy   2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy   2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy   2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy   2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy   2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy   2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy   2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700

Qy   2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy   2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy   2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy   2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy   2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy   3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy   3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy   3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy   3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy   3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy      3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy      3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
```

```
RESULT 5
US-10-010-940-110
; Sequence 110, Application US/10010940
; Publication No. US20030088062A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan Louise
;  APPLICANT:  Jiang Yuqui
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Kalos, Michael
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Retter, Mark
;  APPLICANT:  Solk, John
;  APPLICANT:  Day, Craig
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427D3
;  CURRENT APPLICATION NUMBER: US/10/010,940
;  CURRENT FILING DATE:  2001-12-05
;  NUMBER OF SEQ ID NOS: 575
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 110
;   LENGTH: 3410
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-010-940-110

  Query Match             100.0%;  Score 3409.6;  DB 9;  Length 3410;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy         1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy        61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy       121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy       181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy       241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy       301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy       361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy       421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy       481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy       541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy       601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy       661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy       721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy       781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy       841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy       901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy       961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy      1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy      1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy      1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy      1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy      1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy      1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy      1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy      1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy      1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy       2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy       2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy       2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy       2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCA 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCA 2700

Qy       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAA 3410
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAA 3410

RESULT 6
US-09-745-288-100
; Sequence 100, Application US/09745288
; Patent No. US20010018058A1
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  TITLE OF INVENTION:  COMPOUNDS FOR IMMUNOTHERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
;  FILE REFERENCE:  210121.446D1
;  CURRENT APPLICATION NUMBER: US/09/745,288
;  CURRENT FILING DATE:  2000-12-19
;  NUMBER OF SEQ ID NOS:  101
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 100
;   LENGTH: 3410
;   TYPE: DNA
;   ORGANISM: Homo sapien
```

```
US-09-745-288-100

  Query Match            100.0%;  Score 3409.6;  DB 10;  Length 3410;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
```

```
Qy     2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy     2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy     2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy     2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy     2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy     2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy     2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy     2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy     2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy     2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700

Qy     2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy     2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy     2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy     2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy     2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy     3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy     3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy     3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy     3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy     3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy     3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy     3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
```

```
RESULT 7
US-09-759-143-110
; Sequence 110, Application US/09759143
; Patent No. US20020022248A1
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C23
;  CURRENT APPLICATION NUMBER: US/09/759,143
;  CURRENT FILING DATE:  2001-01-12
;  NUMBER OF SEQ ID NOS:  934
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 110
;   LENGTH: 3410
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-759-143-110

  Query Match             100.0%;  Score 3409.6;  DB 10;  Length 3410;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy    301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy    361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy    421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy    481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy    541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy    601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy    661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy    721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy    781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy    841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy    901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy    961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy   1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy   1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy   1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy   1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy   1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy   1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy   1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy   1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy   1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy   1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy   1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy   1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy   1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy   1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy   1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy   1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy   1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy   2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy   2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy   2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy   2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy   2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy   2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy   2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
```

```
Qy     2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy     2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy     2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy     2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700

Qy     2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy     2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy     2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy     2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy     2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy     3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy     3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy     3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy     3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy     3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy     3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy     3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAA 3410
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAA 3410
```

RESULT 8
US-09-780-669-110
; Sequence 110, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780,669
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 110
; LENGTH: 3410
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-780-669-110

```
  Query Match             100.0%;  Score 3409.6;  DB 10;  Length 3410;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy       61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy      121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy      181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy      241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy      301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy      361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy      421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy      481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy      541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
```

```
Qy    601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy    661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy    721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy    781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy    841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy    901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy    961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy   1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy   1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy   1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy   1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy   1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy   1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy   1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy   1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy   1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy   1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy   1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy   1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy   1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy   1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy   1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy   1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy   1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy   2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy   2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy   2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy   2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy   2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy   2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy   2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy   2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy   2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy   2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy   2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700

Qy   2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy   2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy         2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy         2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy         2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy         3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy         3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy         3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy         3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy         3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy         3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy         3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
```

RESULT 9
US-09-030-606-110
; Sequence 110, Application US/09030606
; Patent No. US20020081580A1
; GENERAL INFORMATION:
;   APPLICANT: Xu, Jiangchun
;   APPLICANT: Dillon, Davin C.
;   TITLE OF INVENTION: COMPOUNDS FOR IMMUNODIAGNOSIS OF PROSTATE CANCER AND METHODS F
;   NUMBER OF SEQUENCES: 224
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: SEED and BERRY LLP
;   STREET: 6300 Columbia Center, 701 Fifth Avenue
;   CITY: Seattle
;   STATE: WA
;   COUNTRY: USA
;   ZIP: 98104
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/030,606
;   FILING DATE: 25-FEB-1998
;   CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Maki, David J.
;   REGISTRATION NUMBER: 31,392
;,   REFERENCE/DOCKET NUMBER: 210121.428C3
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (206) 622-4900
;   TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 110:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 3410 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   ORIGINAL SOURCE:
;   ORGANISM: Homo sapiens
US-09-030-606-110

```
  Query Match             100.0%;  Score 3409.6;  DB 10;  Length 3410;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy            1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy           61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy          121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy          181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy          241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy          301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy          361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy          421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy          481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy          541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy          601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy          661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy          721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy          781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
```

```
Qy    841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy    901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy    961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy   1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy   1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy   1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy   1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy   1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy   1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy   1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy   1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy   1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy   1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy   1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy   1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy   1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy   1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy   1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy   1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy   1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy   2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy   2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy   2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy   2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy   2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy   2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy   2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy   2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy   2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy   2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy   2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700

Qy   2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy   2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy   2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy   2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy   2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy   3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
```

RESULT 10
US-09-822-827-110
; Sequence 110, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 110
;   LENGTH: 3410
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-822-827-110

```
  Query Match             100.0%;  Score 3409.6;  DB 10;  Length 3410;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
```

```
Qy        1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy        1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy        1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy        1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy        1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy        1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy        1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy        1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy        1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy        1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy        2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy        2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy        2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy        2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy        2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy        2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy        2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy        2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy        2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy        2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy        2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCA 2700
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCA 2700

Qy        2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy        2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy        2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy        2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy        2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy        3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy        3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy        3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy        3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy        3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy        3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy        3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
```

```
RESULT 11
US-09-115-453-110
; Sequence 110, Application US/09115453B
; Patent No. US20020090372A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT: Dillon, Davin C.
;  TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND
;  TITLE OF INVENTION: METHODS FOR THEIR USE
;  FILE REFERENCE: 210121.427C4
;  CURRENT APPLICATION NUMBER: US/09/115,453B
;  CURRENT FILING DATE: 1998-07-14
```

```
;  NUMBER OF SEQ ID NOS: 228
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 110
;   LENGTH: 3410
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-115-453-110

  Query Match             100.0%;  Score 3409.6;  DB 10;  Length 3410;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy     2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy     2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy     2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy     2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy     2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy     2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy     2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy     2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy     2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy     2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy     2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700

Qy     2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy     2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy     2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy     2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy     2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy     3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy     3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy     3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy     3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy     3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy     3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy     3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAA 3410
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAA 3410
```

```
RESULT 12
US-09-838-785-1
; Sequence 1, Application US/09838785
; Patent No. US20020009455A1
; GENERAL INFORMATION:
;  APPLICANT: Lau, Ted
;  APPLICANT:  Lin, Rick
;  APPLICANT:  Parkes, Debbie
;  APPLICANT:  Parry, Gordon
;  APPLICANT:  Schneider, Douglas
;  APPLICANT:  Steinbrecher, Renate
;  APPLICANT:  Van Heuit, Pam T
;  APPLICANT:  Wu, John
;  TITLE OF INVENTION: DNA Encoding a No. US20020009455A1el PROST 03
;  FILE REFERENCE: 51831AUSM1
;  CURRENT APPLICATION NUMBER: US/09/838,785
;  CURRENT FILING DATE:  2001-04-20
;  PRIOR APPLICATION NUMBER: 60/200,065
;  PRIOR FILING DATE: 2000-04-27
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 3320
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (282)..(1943)
US-09-838-785-1

  Query Match             96.6%;  Score 3292.4;  DB 10;  Length 3320;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3316;  Conservative    0;  Mismatches    1;  Indels    3;  Gaps    2;

Qy        3 GAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGT 62
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 GAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGT 60

Qy       63 GATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAA 122
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 GATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAA 120

Qy      123 GCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGG 182
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 GCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGG 180

Qy      183 CGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGA 242
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 CGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGA 240

Qy      243 GTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGG 302
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGG 300

Qy        303 TGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTG 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTG 360

Qy        363 GCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGG 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGG 420

Qy        423 GGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCT 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCT 480

Qy        483 GTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGC 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGC 540

Qy        543 CCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCG 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCG 600

Qy        603 GCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCC 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCC 660

Qy        663 TGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGC 722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGC 720

Qy        723 TCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCA 782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCA 780

Qy        783 TGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTG 842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTG 840

Qy        843 CCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCT 902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCT 900

Qy        903 TCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCG 962
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCG 960

Qy        963 AGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCC 1022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCC 1020

Qy       1023 GCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCA 1082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCA 1080

Qy       1083 TGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGA 1142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGA 1140

Qy       1143 CCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAG 1202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAG 1200

Qy       1203 CTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGG 1262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGG 1260

Qy       1263 GGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGC 1322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGC 1320

Qy       1323 AGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCG 1382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCG 1380

Qy       1383 GTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGT 1442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGT 1440

Qy       1443 TCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGA 1502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGA 1500

Qy       1503 AGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCC 1562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCC 1560

Qy       1563 TGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGG 1622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGG 1620

Qy       1623 GTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTG 1682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTG 1680

Qy       1683 ATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGG 1742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 ATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGG 1740

Qy       1743 GCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCAT 1802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCAT 1800

Qy       1803 CCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTG 1862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTG 1860

Qy       1863 CCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCG 1922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCG 1920

Qy       1923 ACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCA 1982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 ACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCA 1980

Qy       1983 CTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTT 2042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTT 2040

Qy       2043 CTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGC 2102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 CTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGC 2100

Qy       2103 TGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGAC 2162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGAC 2160

Qy       2163 TGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCAT 2222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 TGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCAT 2220

Qy       2223 GCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCT 2282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 GCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCT 2280

Qy       2283 CCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGT 2342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGT 2340

Qy       2343 TTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTT 2402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 TTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTT 2400

Qy       2403 TCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAGA 2460
              |||||||||||||||||||||||||||||||||||||||  | |||||||||||||||||
Db       2401 TCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGAAAGTTATTTGTAGGGGAAGA 2460
```

```
Qy       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCA 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCA 2700

Qy       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy       3181 GAGGTCTTATCTCTCA-GGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTAT 3239
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GAGGTCTTATCTCTCAGGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTAT 3240

Qy       3240 TTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGAC 3299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGAC 3300

Qy       3300 AAAATTAAAGGCTTTCTTAT 3319
              ||||||||||||||||||||
Db       3301 AAAATTAAAGGCTTTCTTAT 3320
```

RESULT 13
US-10-012-896-704
; Sequence 704, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Wantanabe, Yoshihiro
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012,896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 704
;   LENGTH: 4034
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-012-896-704

```
  Query Match             75.8%;  Score 2585.4;  DB 9;  Length 4034;
  Best Local Similarity   82.4%;  Pred. No. 0;
  Matches 3325;  Conservative    0;  Mismatches    1;  Indels  709;  Gaps    3;

Qy          4 AACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTG 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTG 60

Qy         64 ATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAG 120

Qy        124 CTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGC 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGC 180

Qy        184 GGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAG 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAG 240

Qy        244 TGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGT 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGT 300

Qy        304 GAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGG 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGG 360

Qy        364 CCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGG 420

Qy        424 GGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTG 480

Qy        484 TGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC 540

Qy        544 CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGG 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      541 CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGG 600

Qy      604 CTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCT 663
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 CTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCT 660

Qy      664 GGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT 723
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 GGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT 720

Qy      724 CTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCAT 783
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 CTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCAT 780

Qy      784 GATCAGTCTTGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGC 843
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 GATCAGTCTTGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGC 840

Qy      844 CCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTT 903
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 CCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTT 900

Qy      904 CCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGA 963
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 CCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGA 960

Qy      964 GCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCG 1023
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 GCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCG 1020

Qy     1024 CTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCAT 1083
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 CTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCAT 1080

Qy     1084 GCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGAC 1143
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 GCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGAC 1140

Qy     1144 CTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGC 1203
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 CTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGC 1200

Qy     1204 TGAGCCGGGCACCGAGGCCCGGAGACACTATGATGA------------------------ 1239
            ||||||||||||||||||||||||||||||||||||
Db     1201 TGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGTAAGGCCTTGGCAGCCAGCAG 1260

Qy     1240 ------------------------------------------------------------ 1239

Db     1261 AGGCTGGTGTGGGAGCCGCCCACCAGAGACGACACTCGGGGCTGTGTCTGGGCTGGTGCC 1320

Qy     1240 ------------------------------------------------------------ 1239

Db     1321 TCTCCATCCTGGCCCCGACTTCTCTGTCAGGAAAGTGGGGATGGACCCCATCTGCATACA 1380

Qy     1240 ------------------------------------------------------------ 1239

Db     1381 CGGCTTCTCATGGGTGTGGAACATCTCTGCTTGCGGTTTCAGGAAGGCCTCTGGCTGCTC 1440

Qy     1240 ------------------------------------------------------------ 1239

Db     1441 TAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGACAGAAGGAAAGGCGGAGCTTATTCAA 1500

Qy     1240 ------------------------------------------------------------ 1239

Db     1501 AGTCTAGAGGGAGTGGAGGAGTTAAGGCTGGATTTCAGATCTGCCTGGTTCCAGCCGCAG 1560

Qy     1240 ------------------------------------------------------------ 1239

Db     1561 TGTGCCCTCTGCTCCCCCAACGACTTTCCAAATAATCTCACCAGCGCCTTCCAGCTCAGG 1620

Qy     1240 ------------------------------------------------------------ 1239

Db     1621 CGTCCTAGAAGCGTCTTGAAGCCTATGGCCAGCTGTCTTTGTGTTCCCTCTCACCCGCCT 1680

Qy     1240 ------------------------------------------------------------ 1239

Db     1681 GTCCTCACAGCTGAGACTCCCAGGAAACCTTCAGACTACCTTCCTCTGCCTTCAGCAAGG 1740

Qy     1240 ------------------------------------------------------------ 1239

Db     1741 GGCGTTGCCCACATTCTCTGAGGGTCAGTGGAAGAACCTAGACTCCCATTGCTAGAGGTA 1800

Qy     1240 ------------------------------------------------------------ 1239

Db     1801 GAAAGGGGAAGGGTGCTGGGGAGCAGGGCTGGTCCACAGCAGGTCTCGTGCAGCAGGTAC 1860

Qy     1240 ------------------------------------------------------------ 1239

Db     1861 CTGTGGTTCCGCCTTCTCATCTCCCTGAGACTGCTCCGACCCTTCCCTCCCAGGCTCTGT 1920

Qy     1240 ----------------------AGGCGTTCGGATGGCAGCCTGGGGCTGTTCCTGCAGT 1277
                                  |||||||||||||||||||||||||||||||||||||
Db     1921 CTGATGGCCCCTCTCCCTCTGCAGGCGTTCGGATGGCAGCCTGGGGCTGTTCCTGCAGT 1980

Qy     1278 GCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTC 1337
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1981 GCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTC 2040

Qy     1338 GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT 1397
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2041 GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT 2100

Qy     1398 CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCC 1457
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2101 CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCC 2160

Qy     1458 TGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGC 1517
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2161 TGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGC 2220

Qy     1518 CCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCC 1577
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2221 CCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCC 2280

Qy     1578 TGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTG 1637
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2281 TGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTG 2340

Qy     1638 GCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTG 1697
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2341 GCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTG 2400

Qy     1698 TGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACC 1757
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2401 TGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACC 2460

Qy     1758 TCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCT 1817
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2461 TCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCT 2520

Qy     1818 CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTC 1877
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2521 CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTC 2580

Qy     1878 TGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT 1937
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2581 TGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT 2640

Qy     1938 CAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCC 1997
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2641 CAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCC 2700

Qy     1998 CCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAG 2057
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2701 CCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAG 2760
```

```
Qy       2058 TAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGC 2117
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGC 2820

Qy       2118 TGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAG 2177
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAG 2880

Qy       2178 GGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGG 2237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGG 2940

Qy       2238 GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACA 2297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACA 3000

Qy       2298 CCTAGAGAAGGGTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGC 2357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCTAGAGAAGGGTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGC 3060

Qy       2358 CCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACAC 2417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACAC 3120

Qy       2418 TCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAAC 2475
              ||||||||||||||||||||||||   |  ||||||||||||||||||||||||||||||
Db       3121 TCCTCCATGGGATTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAAC 3180

Qy       2476 ACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCT 2535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 ACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCT 3240

Qy       2536 TACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCAT 2595
              |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db       3241 TACCTTTTATCAGGATGT-GCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCAT 3299

Qy       2596 TTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGT 2655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3300 TTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGT 3359

Qy       2656 GTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAG 2715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3360 GTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAG 3419

Qy       2716 CTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAAT 2775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3420 CTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAAT 3479

Qy       2776 GCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTAC 2835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3480 GCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTAC 3539

Qy       2836 CCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT 2895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3540 CCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT 3599

Qy       2896 AACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTC 2955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3600 AACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTC 3659

Qy       2956 TCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTA 3015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3660 TCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTA 3719

Qy       3016 CCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCA 3075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3720 CCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCA 3779

Qy       3076 CAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGG 3135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3780 CAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGG 3839

Qy       3136 GAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTC 3195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3840 GAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTC 3899

Qy       3196 AGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATAT 3255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3900 AGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATAT 3959

Qy       3256 TTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTC 3315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3960 TTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTC 4019

Qy       3316 TTATATGTTTAAAAA 3330
              |||||||||||||||
Db       4020 TTATATGTTTAAAAA 4034
```

```
RESULT 14
US-09-895-793-704
; Sequence 704, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.534C2
;  CURRENT APPLICATION NUMBER: US/09/895,793
;  CURRENT FILING DATE:  2001-06-29
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 704
;   LENGTH: 4034
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-895-793-704

  Query Match            75.8%;  Score 2585.4;  DB 9;  Length 4034;
  Best Local Similarity   82.4%;  Pred. No. 0;
  Matches 3325;  Conservative    0;  Mismatches    1;  Indels  709;  Gaps    3;

Qy          4 AACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTG 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTG 60

Qy         64 ATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAG 120

Qy        124 CTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGC 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGC 180

Qy        184 GGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAG 243
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAG 240

Qy        244 TGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGT 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGT 300

Qy        304 GAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGG 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGG 360

Qy        364 CCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGG 420

Qy        424 GGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTG 480

Qy        484 TGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC 540

Qy        544 CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGG 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGG 600

Qy        604 CTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCT 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCT 660

Qy        664 GGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT 720

Qy        724 CTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCAT 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCAT 780

Qy        784 GATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGC 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGC 840

Qy        844 CCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTT 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTT 900

Qy        904 CCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGA 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGA 960

Qy        964 GCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCG 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCG 1020

Qy       1024 CTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCAT 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCAT 1080

Qy       1084 GCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGAC 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGAC 1140

Qy       1144 CTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGC 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGC 1200

Qy       1204 TGAGCCGGGCACCGAGGCCCGGAGACACTATGATGA------------------------ 1239
              ||||||||||||||||||||||||||||||||||||
Db       1201 TGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGTAAGGCCTTGGCAGCCAGCAG 1260

Qy       1240 ------------------------------------------------------------ 1239

Db       1261 AGGCTGGTGTGGGAGCCGCCCACCAGAGACGACACTCGGGGCTGTGTCTGGGCTGGTGCC 1320

Qy       1240 ------------------------------------------------------------ 1239

Db       1321 TCTCCATCCTGGCCCCGACTTCTCTGTCAGGAAAGTGGGGATGGACCCCATCTGCATACA 1380

Qy       1240 ------------------------------------------------------------ 1239

Db       1381 CGGCTTCTCATGGGTGTGGAACATCTCTGCTTGCGGTTTCAGGAAGGCCTCTGGCTGCTC 1440

Qy       1240 ------------------------------------------------------------ 1239

Db       1441 TAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGACAGAAGGAAAGGCGGAGCTTATTCAA 1500

Qy       1240 ------------------------------------------------------------ 1239

Db       1501 AGTCTAGAGGGAGTGGAGGAGTTAAGGCTGGATTTCAGATCTGCCTGGTTCCAGCCGCAG 1560

Qy       1240 ------------------------------------------------------------ 1239

Db       1561 TGTGCCCTCTGCTCCCCCAACGACTTTCCAAATAATCTCACCAGCGCCTTCCAGCTCAGG 1620

Qy       1240 ------------------------------------------------------------ 1239

Db       1621 CGTCCTAGAAGCGTCTTGAAGCCTATGGCCAGCTGTCTTTGTGTTCCCTCTCACCCGCCT 1680

Qy       1240 ------------------------------------------------------------ 1239

Db       1681 GTCCTCACAGCTGAGACTCCCAGGAAACCTTCAGACTACCTTCCTCTGCCTTCAGCAAGG 1740

Qy       1240 ------------------------------------------------------------ 1239

Db       1741 GGCGTTGCCCACATTCTCTGAGGGTCAGTGGAAGAACCTAGACTCCCATTGCTAGAGGTA 1800

Qy       1240 ------------------------------------------------------------ 1239

Db       1801 GAAAGGGGAAGGGTGCTGGGGAGCAGGGCTGGTCCACAGCAGGTCTCGTGCAGCAGGTAC 1860

Qy       1240 ------------------------------------------------------------ 1239

Db       1861 CTGTGGTTCCGCCTTCTCATCTCCCTGAGACTGCTCCGACCCTTCCCTCCCAGGCTCTGT 1920

Qy       1240 ----------------------AGGCGTTCGGATGGCAGCCTGGGGCTGTTCCTGCAGT 1277
                                    ||||||||||||||||||||||||||||||||||||||
Db       1921 CTGATGGCCCCTCTCCCTCTGCAGGCGTTCGGATGGCAGCCTGGGGCTGTTCCTGCAGT 1980

Qy       1278 GCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTC 1337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTC 2040

Qy       1338 GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT 1397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT 2100

Qy       1398 CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCC 1457
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCC 2160

Qy       1458 TGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGC 1517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 TGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGC 2220

Qy       1518 CCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCC 1577
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCC 2280

Qy       1578 TGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTG 1637
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 TGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTG 2340

Qy       1638 GCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTG 1697
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTG 2400
```

```
Qy       1698 TGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACC 1757
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACC 2460

Qy       1758 TCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCT 1817
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 TCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCT 2520

Qy       1818 CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTC 1877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTC 2580

Qy       1878 TGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT 1937
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 TGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT 2640

Qy       1938 CAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCC 1997
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCC 2700

Qy       1998 CCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAG 2057
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 CCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAG 2760

Qy       2058 TAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2117
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2820

Qy       2118 TGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAG 2177
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAG 2880

Qy       2178 GGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGG 2237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGG 2940

Qy       2238 GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACA 2297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACA 3000

Qy       2298 CCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGC 2357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGC 3060

Qy       2358 CCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACAC 2417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACAC 3120

Qy       2418 TCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAAC 2475
              ||||||||||||||||||||||||  | ||||||||||||||||||||||||||||||||
Db       3121 TCCTCCATGGGATTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAAC 3180

Qy       2476 ACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCT 2535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 ACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCT 3240

Qy       2536 TACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCAT 2595
              |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db       3241 TACCTTTTATCAGGATGT-GCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCAT 3299

Qy       2596 TTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGT 2655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3300 TTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGT 3359

Qy       2656 GTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAG 2715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3360 GTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAG 3419

Qy       2716 CTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAAT 2775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3420 CTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAAT 3479

Qy       2776 GCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTAC 2835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3480 GCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTAC 3539

Qy       2836 CCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT 2895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3540 CCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT 3599

Qy       2896 AACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTC 2955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3600 AACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTC 3659

Qy       2956 TCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTA 3015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3660 TCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTA 3719

Qy       3016 CCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCA 3075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3720 CCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCA 3779

Qy       3076 CAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGG 3135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3780 CAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGG 3839

Qy       3136 GAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTC 3195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3840 GAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTC 3899

Qy       3196 AGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATAT 3255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3900 AGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATAT 3959

Qy       3256 TTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTC 3315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3960 TTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTC 4019

Qy       3316 TTATATGTTTAAAAA 3330
              |||||||||||||||
Db       4020 TTATATGTTTAAAAA 4034
```

```
RESULT 15
US-09-895-814-704
; Sequence 704, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C26
;  CURRENT APPLICATION NUMBER: US/09/895,814
;  CURRENT FILING DATE:  2001-06-29
```

```
;  NUMBER OF SEQ ID NOS: 990
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 704
;   LENGTH: 4034
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-895-814-704

  Query Match            75.8%;  Score 2585.4;  DB 9;  Length 4034;
  Best Local Similarity  82.4%;  Pred. No. 0;
  Matches 3325;  Conservative    0;  Mismatches    1;  Indels  709;  Gaps    3;

Qy          4 AACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTG 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTG 60

Qy         64 ATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAG 120

Qy        124 CTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGC 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGC 180

Qy        184 GGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAG 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAG 240

Qy        244 TGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGT 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGT 300

Qy        304 GAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGG 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGG 360

Qy        364 CCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGG 420

Qy        424 GGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTG 480

Qy        484 TGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC 540

Qy        544 CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGG 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGG 600

Qy        604 CTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCT 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCT 660

Qy        664 GGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT 720

Qy        724 CTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCAT 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCAT 780

Qy        784 GATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGC 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGC 840

Qy        844 CCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTT 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTT 900

Qy        904 CCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGA 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGA 960

Qy        964 GCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCG 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCG 1020

Qy       1024 CTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCAT 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCAT 1080

Qy       1084 GCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGAC 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGAC 1140

Qy       1144 CTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGC 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGC 1200

Qy       1204 TGAGCCGGGCACCGAGGCCCGGAGACACTATGATGA------------------------ 1239
              ||||||||||||||||||||||||||||||||||||
Db       1201 TGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGTAAGGCCTTGGCAGCCAGCAG 1260

Qy       1240 ------------------------------------------------------------ 1239

Db       1261 AGGCTGGTGTGGGAGCCGCCCACCAGAGACGACACTCGGGGCTGTGTCTGGGCTGGTGCC 1320

Qy       1240 ------------------------------------------------------------ 1239

Db       1321 TCTCCATCCTGGCCCCGACTTCTCTGTCAGGAAAGTGGGGATGGACCCCATCTGCATACA 1380

Qy       1240 ------------------------------------------------------------ 1239

Db       1381 CGGCTTCTCATGGGTGTGGAACATCTCTGCTTGCGGTTTCAGGAAGGCCTCTGGCTGCTC 1440

Qy       1240 ------------------------------------------------------------ 1239

Db       1441 TAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGACAGAAGGAAAGGCGGAGCTTATTCAA 1500

Qy       1240 ------------------------------------------------------------ 1239

Db       1501 AGTCTAGAGGGAGTGGAGGAGTTAAGGCTGGATTTCAGATCTGCCTGGTTCCAGCCGCAG 1560

Qy       1240 ------------------------------------------------------------ 1239

Db       1561 TGTGCCCTCTGCTCCCCCAACGACTTTCCAAATAATCTCACCAGCGCCTTCCAGCTCAGG 1620

Qy       1240 ------------------------------------------------------------ 1239

Db       1621 CGTCCTAGAAGCGTCTTGAAGCCTATGGCCAGCTGTCTTTGTGTTCCCTCTCACCCGCCT 1680

Qy       1240 ------------------------------------------------------------ 1239

Db       1681 GTCCTCACAGCTGAGACTCCCAGGAAACCTTCAGACTACCTTCCTCTGCCTTCAGCAAGG 1740

Qy       1240 ------------------------------------------------------------ 1239

Db       1741 GGCGTTGCCCACATTCTCTGAGGGTCAGTGGAAGAACCTAGACTCCCATTGCTAGAGGTA 1800

Qy       1240 ------------------------------------------------------------ 1239

Db       1801 GAAAGGGGAAGGGTGCTGGGGAGCAGGGCTGGTCCACAGCAGGTCTCGTGCAGCAGGTAC 1860

Qy       1240 ------------------------------------------------------------ 1239

Db       1861 CTGTGGTTCCGCCTTCTCATCTCCCTGAGACTGCTCCGACCCTTCCCTCCCAGGCTCTGT 1920

Qy       1240 ----------------------AGGCGTTCGGATGGCAGCCTGGGGCTGTTCCTGCAGT 1277
                                    ||||||||||||||||||||||||||||||||||||||
Db       1921 CTGATGGCCCCTCTCCCTCTGCAGGCGTTCGGATGGCAGCCTGGGGCTGTTCCTGCAGT 1980

Qy       1278 GCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTC 1337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     1981 GCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTC 2040

Qy     1338 GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT 1397
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2041 GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT 2100

Qy     1398 CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCC 1457
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2101 CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCC 2160

Qy     1458 TGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGC 1517
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2161 TGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGC 2220

Qy     1518 CCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCC 1577
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2221 CCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCC 2280

Qy     1578 TGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTG 1637
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2281 TGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTG 2340

Qy     1638 GCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTG 1697
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2341 GCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTG 2400

Qy     1698 TGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACC 1757
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2401 TGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACC 2460

Qy     1758 TCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCT 1817
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2461 TCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCT 2520

Qy     1818 CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTC 1877
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2521 CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTC 2580

Qy     1878 TGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT 1937
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2581 TGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT 2640

Qy     1938 CAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCC 1997
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2641 CAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCC 2700

Qy     1998 CCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAG 2057
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2701 CCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAG 2760

Qy     2058 TAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2117
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2761 TAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2820

Qy     2118 TGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAG 2177
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2821 TGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAG 2880

Qy     2178 GGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGG 2237
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2881 GGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGG 2940

Qy     2238 GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACA 2297
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2941 GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACA 3000

Qy     2298 CCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGC 2357
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3001 CCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGC 3060

Qy     2358 CCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACAC 2417
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3061 CCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACAC 3120

Qy     2418 TCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAAC 2475
            ||||||||||||||||||||||||  | ||||||||||||||||||||||||||||||||
Db     3121 TCCTCCATGGGATTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAAC 3180

Qy     2476 ACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTGCTGATCCACCCCCCTCT 2535
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3181 ACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTGCTGATCCACCCCCCTCT 3240

Qy     2536 TACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCAT 2595
            |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db     3241 TACCTTTTATCAGGATGT-GCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCAT 3299

Qy     2596 TTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGT 2655
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3300 TTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGT 3359

Qy     2656 GTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAG 2715
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3360 GTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAG 3419

Qy     2716 CTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAAT 2775
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3420 CTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAAT 3479

Qy     2776 GCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTAC 2835
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3480 GCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTAC 3539

Qy     2836 CCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT 2895
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3540 CCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT 3599

Qy     2896 AACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTC 2955
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3600 AACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTC 3659

Qy     2956 TCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTA 3015
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3660 TCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTA 3719

Qy     3016 CCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCA 3075
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3720 CCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCA 3779

Qy     3076 CAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGG 3135
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3780 CAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGG 3839

Qy     3136 GAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTC 3195
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3840 GAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTC 3899

Qy     3196 AGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATAT 3255
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3900 AGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATAT 3959

Qy     3256 TTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTC 3315
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3960 TTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTC 4019

Qy     3316 TTATATGTTTAAAAA 3330
            |||||||||||||||
Db     4020 TTATATGTTTAAAAA 4034
```

```
RESULT 16
US-09-759-143-704
; Sequence 704, Application US/09759143
; Patent No. US20020022248A1
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
```

```
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C23
;  CURRENT APPLICATION NUMBER: US/09/759,143
;  CURRENT FILING DATE:  2001-01-12
;  NUMBER OF SEQ ID NOS:  934
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 704
;   LENGTH: 4034
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-759-143-704

  Query Match            75.8%;  Score 2585.4;  DB 10;  Length 4034;
  Best Local Similarity  82.4%;  Pred. No. 0;
  Matches 3325;  Conservative    0;  Mismatches    1;  Indels  709;  Gaps    3;

Qy          4 AACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTG 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTG 60

Qy         64 ATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAG 120

Qy        124 CTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGC 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGC 180

Qy        184 GGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAG 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAG 240

Qy        244 TGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGT 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGT 300

Qy        304 GAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGG 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGG 360

Qy        364 CCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGG 420

Qy        424 GGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTG 480

Qy        484 TGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC 540

Qy        544 CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGG 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGG 600

Qy        604 CTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCT 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCT 660

Qy        664 GGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT 720

Qy        724 CTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCAT 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCAT 780

Qy        784 GATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGC 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGC 840

Qy        844 CCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTT 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTT 900

Qy        904 CCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGA 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGA 960

Qy        964 GCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCG 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCG 1020

Qy       1024 CTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCAT 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCAT 1080

Qy       1084 GCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGAC 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGAC 1140

Qy       1144 CTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGC 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGC 1200

Qy       1204 TGAGCCGGGCACCGAGGCCCGGAGACACTATGATGA------------------------ 1239
              ||||||||||||||||||||||||||||||||||||
Db       1201 TGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGTAAGGCCTTGGCAGCCAGCAG 1260

Qy       1240 ------------------------------------------------------------ 1239

Db       1261 AGGCTGGTGTGGGAGCCGCCCACCAGAGACGACACTCGGGGCTGTGTCTGGGCTGGTGCC 1320

Qy       1240 ------------------------------------------------------------ 1239

Db       1321 TCTCCATCCTGGCCCCGACTTCTCTGTCAGGAAAGTGGGGATGGACCCCATCTGCATACA 1380

Qy       1240 ------------------------------------------------------------ 1239

Db       1381 CGGCTTCTCATGGGTGTGGAACATCTCTGCTTGCGGTTTCAGGAAGGCCTCTGGCTGCTC 1440

Qy       1240 ------------------------------------------------------------ 1239

Db       1441 TAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGACAGAAGGAAAGGCGGAGCTTATTCAA 1500

Qy       1240 ------------------------------------------------------------ 1239

Db       1501 AGTCTAGAGGGAGTGGAGGAGTTAAGGCTGGATTTCAGATCTGCCTGGTTCCAGCCGCAG 1560

Qy       1240 ------------------------------------------------------------ 1239

Db       1561 TGTGCCCTCTGCTCCCCCAACGACTTTCCAAATAATCTCACCAGCGCCTTCCAGCTCAGG 1620

Qy       1240 ------------------------------------------------------------ 1239

Db       1621 CGTCCTAGAAGCGTCTTGAAGCCTATGGCCAGCTGTCTTTGTGTTCCCTCTCACCCGCCT 1680

Qy       1240 ------------------------------------------------------------ 1239
```

```
Db     1681 GTCCTCACAGCTGAGACTCCCAGGAAACCTTCAGACTACCTTCCTCTGCCTTCAGCAAGG 1740

Qy     1240 ------------------------------------------------------------ 1239

Db     1741 GGCGTTGCCCACATTCTCTGAGGGTCAGTGGAAGAACCTAGACTCCCATTGCTAGAGGTA 1800

Qy     1240 ------------------------------------------------------------ 1239

Db     1801 GAAAGGGGAAGGGTGCTGGGGAGCAGGGCTGGTCCACAGCAGGTCTCGTGCAGCAGGTAC 1860

Qy     1240 ------------------------------------------------------------ 1239

Db     1861 CTGTGGTTCCGCCTTCTCATCTCCCTGAGACTGCTCCGACCCTTCCCTCCCAGGCTCTGT 1920

Qy     1240 ----------------------AGGCGTTCGGATGGCAGCCTGGGGCTGTTCCTGCAGT 1277
                                  ||||||||||||||||||||||||||||||||||||||
Db     1921 CTGATGGCCCCTCTCCCTCTGCAGGCGTTCGGATGGCAGCCTGGGGCTGTTCCTGCAGT 1980

Qy     1278 GCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTC 1337
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1981 GCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTC 2040

Qy     1338 GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT 1397
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2041 GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT 2100

Qy     1398 CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCC 1457
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2101 CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCC 2160

Qy     1458 TGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGC 1517
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2161 TGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGC 2220

Qy     1518 CCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCC 1577
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2221 CCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCC 2280

Qy     1578 TGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTG 1637
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2281 TGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTG 2340

Qy     1638 GCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTG 1697
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2341 GCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTG 2400

Qy     1698 TGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACC 1757
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2401 TGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACC 2460

Qy     1758 TCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCT 1817
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2461 TCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCT 2520

Qy     1818 CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTC 1877
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2521 CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTC 2580

Qy     1878 TGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT 1937
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2581 TGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT 2640

Qy     1938 CAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGTCCCAGCTCC 1997
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2641 CAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGTCCCAGCTCC 2700

Qy     1998 CCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAG 2057
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2701 CCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAG 2760

Qy     2058 TAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2117
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2761 TAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2820

Qy     2118 TGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAG 2177
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2821 TGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAG 2880

Qy     2178 GGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGG 2237
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2881 GGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGG 2940

Qy     2238 GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACA 2297
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2941 GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACA 3000

Qy     2298 CCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGC 2357
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3001 CCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGC 3060

Qy     2358 CCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACAC 2417
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3061 CCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACAC 3120

Qy     2418 TCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAAC 2475
            ||||||||||||||||||||||||  | ||||||||||||||||||||||||||||||||
Db     3121 TCCTCCATGGGATTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAAC 3180

Qy     2476 ACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCT 2535
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3181 ACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCT 3240

Qy     2536 TACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCAT 2595
            |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db     3241 TACCTTTTATCAGGATGT-GCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCAT 3299

Qy     2596 TTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGT 2655
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3300 TTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGT 3359

Qy     2656 GTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAG 2715
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3360 GTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAG 3419

Qy     2716 CTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAAT 2775
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3420 CTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAAT 3479

Qy     2776 GCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTAC 2835
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3480 GCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTAC 3539

Qy     2836 CCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT 2895
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3540 CCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT 3599

Qy     2896 AACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTC 2955
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3600 AACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTC 3659

Qy     2956 TCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTA 3015
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3660 TCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTA 3719

Qy     3016 CCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCA 3075
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3720 CCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCA 3779

Qy     3076 CAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGG 3135
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3780 CAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGG 3839

Qy     3136 GAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTC 3195
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3840 GAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTC 3899
```

```
Qy       3196 AGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATAT 3255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3900 AGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATAT 3959

Qy       3256 TTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTC 3315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3960 TTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTC 4019

Qy       3316 TTATATGTTTAAAAA 3330
              |||||||||||||||
Db       4020 TTATATGTTTAAAAA 4034

RESULT 17
US-09-780-669-704
; Sequence 704, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C24
;  CURRENT APPLICATION NUMBER: US/09/780,669
;  CURRENT FILING DATE:  2001-02-09
;  NUMBER OF SEQ ID NOS:  943
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 704
;   LENGTH: 4034
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-780-669-704

  Query Match             75.8%;  Score 2585.4;  DB 10;  Length 4034;
  Best Local Similarity   82.4%;  Pred. No. 0;
  Matches 3325;  Conservative    0;  Mismatches    1;  Indels  709;  Gaps    3;

Qy          4 AACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTG 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTG 60

Qy         64 ATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAG 120

Qy        124 CTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGC 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGC 180

Qy        184 GGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAG 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAG 240

Qy        244 TGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGT 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGT 300

Qy        304 GAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGG 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGG 360

Qy        364 CCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGG 420

Qy        424 GGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTG 480

Qy        484 TGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCGGCCGCCGGCC 540

Qy        544 CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGG 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGG 600

Qy        604 CTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCT 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCT 660

Qy        664 GGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT 720

Qy        724 CTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCAT 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCAT 780

Qy        784 GATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGC 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGC 840

Qy        844 CCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTT 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTT 900

Qy        904 CCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGA 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGA 960

Qy        964 GCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCG 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCG 1020

Qy       1024 CTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCAT 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCAT 1080

Qy       1084 GCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGAC 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGAC 1140

Qy       1144 CTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGC 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGC 1200

Qy       1204 TGAGCCGGGCACCGAGGCCCGGAGACACTATGATGA------------------------ 1239
              ||||||||||||||||||||||||||||||||||||
Db       1201 TGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGTAAGGCCTTGGCAGCCAGCAG 1260

Qy       1240 ------------------------------------------------------------ 1239

Db       1261 AGGCTGGTGTGGGAGCCGCCCACCAGAGACGACACTCGGGGCTGTGTCTGGGCTGGTGCC 1320

Qy       1240 ------------------------------------------------------------ 1239

Db       1321 TCTCCATCCTGGCCCCGACTTCTCTGTCAGGAAAGTGGGGATGGACCCCATCTGCATACA 1380
```

```
Qy       1240 ------------------------------------------------------------ 1239

Db       1381 CGGCTTCTCATGGGTGTGGAACATCTCTGCTTGCGGTTTCAGGAAGGCCTCTGGCTGCTC 1440

Qy       1240 ------------------------------------------------------------ 1239

Db       1441 TAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGACAGAAGGAAAGGCGGAGCTTATTCAA 1500

Qy       1240 ------------------------------------------------------------ 1239

Db       1501 AGTCTAGAGGGAGTGGAGGAGTTAAGGCTGGATTTCAGATCTGCCTGGTTCCAGCCGCAG 1560

Qy       1240 ------------------------------------------------------------ 1239

Db       1561 TGTGCCCTCTGCTCCCCCAACGACTTTCCAAATAATCTCACCAGCGCCTTCCAGCTCAGG 1620

Qy       1240 ------------------------------------------------------------ 1239

Db       1621 CGTCCTAGAAGCGTCTTGAAGCCTATGGCCAGCTGTCTTTGTGTTCCCTCTCACCCGCCT 1680

Qy       1240 ------------------------------------------------------------ 1239

Db       1681 GTCCTCACAGCTGAGACTCCCAGGAAACCTTCAGACTACCTTCCTCTGCCTTCAGCAAGG 1740

Qy       1240 ------------------------------------------------------------ 1239

Db       1741 GGCGTTGCCCACATTCTCTGAGGGTCAGTGGAAGAACCTAGACTCCCATTGCTAGAGGTA 1800

Qy       1240 ------------------------------------------------------------ 1239

Db       1801 GAAAGGGGAAGGGTGCTGGGGAGCAGGGCTGGTCCACAGCAGGTCTCGTGCAGCAGGTAC 1860

Qy       1240 ------------------------------------------------------------ 1239

Db       1861 CTGTGGTTCCGCCTTCTCATCTCCCTGAGACTGCTCCGACCCTTCCCTCCCAGGCTCTGT 1920

Qy       1240 ----------------------AGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGT 1277
                                    ||||||||||||||||||||||||||||||||||||||
Db       1921 CTGATGGCCCCTCTCCCTCTGCAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGT 1980

Qy       1278 GCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTC 1337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTC 2040

Qy       1338 GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT 1397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT 2100

Qy       1398 CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCC 1457
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCC 2160

Qy       1458 TGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGC 1517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 TGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGC 2220

Qy       1518 CCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCC 1577
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCC 2280

Qy       1578 TGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTG 1637
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 TGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTG 2340

Qy       1638 GCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTG 1697
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTG 2400

Qy       1698 TGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACC 1757
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACC 2460

Qy       1758 TCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCT 1817
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 TCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCT 2520

Qy       1818 CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTC 1877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTC 2580

Qy       1878 TGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT 1937
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 TGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT 2640

Qy       1938 CAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCC 1997
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCC 2700

Qy       1998 CCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAG 2057
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 CCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAG 2760

Qy       2058 TAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2117
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2820

Qy       2118 TGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAG 2177
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAG 2880

Qy       2178 GGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGG 2237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGG 2940

Qy       2238 GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACA 2297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACA 3000

Qy       2298 CCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGC 2357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGC 3060

Qy       2358 CCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACAC 2417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACAC 3120

Qy       2418 TCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAAC 2475
              ||||||||||||||||||||||||  | ||||||||||||||||||||||||||||||||
Db       3121 TCCTCCATGGGATTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAAC 3180

Qy       2476 ACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCT 2535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 ACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCT 3240

Qy       2536 TACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCAT 2595
              |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db       3241 TACCTTTTATCAGGATGT-GCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCAT 3299

Qy       2596 TTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGT 2655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3300 TTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGT 3359

Qy       2656 GTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAG 2715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3360 GTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAG 3419

Qy       2716 CTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAAT 2775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3420 CTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAAT 3479

Qy       2776 GCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTAC 2835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3480 GCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTAC 3539

Qy       2836 CCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT 2895
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3540 CCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT 3599

Qy        2896 AACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCACTTCCACTCCCCTCTACTCTC 2955
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3600 AACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCACTTCCACTCCCCTCTACTCTC 3659

Qy        2956 TCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTA 3015
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3660 TCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTA 3719

Qy        3016 CCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCA 3075
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3720 CCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCA 3779

Qy        3076 CAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGG 3135
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3780 CAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGG 3839

Qy        3136 GAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTC 3195
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3840 GAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTC 3899

Qy        3196 AGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATAT 3255
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3900 AGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATAT 3959

Qy        3256 TTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTC 3315
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3960 TTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTC 4019

Qy        3316 TTATATGTTTAAAAA 3330
               |||||||||||||||
Db        4020 TTATATGTTTAAAAA 4034
```

```
RESULT 18
US-09-822-827-704
; Sequence 704, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.534C1
;  CURRENT APPLICATION NUMBER: US/09/822,827
;  CURRENT FILING DATE:  2001-03-28
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 704
;   LENGTH: 4034
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-822-827-704

  Query Match             75.8%;  Score 2585.4;  DB 10;  Length 4034;
  Best Local Similarity   82.4%;  Pred. No. 0;
  Matches 3325;  Conservative    0;  Mismatches    1;  Indels  709;  Gaps    3;

Qy           4 AACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTG 63
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 AACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTG 60

Qy          64 ATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAG 123
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 ATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAG 120

Qy         124 CTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGC 183
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGC 180

Qy         184 GGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAG 243
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 GGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAG 240

Qy         244 TGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGT 303
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 TGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGT 300

Qy         304 GAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGG 363
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGG 360

Qy         364 CCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGG 423
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 CCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGG 420

Qy         424 GGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTG 483
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 GGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTG 480

Qy         484 TGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC 543
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 TGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC 540

Qy         544 CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGG 603
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGG 600

Qy         604 CTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCT 663
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 CTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCT 660

Qy         664 GGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT 723
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 GGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT 720

Qy         724 CTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCAT 783
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 CTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCAT 780

Qy         784 GATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGC 843
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 GATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGC 840

Qy         844 CCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTT 903
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 CCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTT 900

Qy         904 CCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGA 963
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 CCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGA 960

Qy         964 GCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCG 1023
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 GCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCG 1020

Qy        1024 CTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCAT 1083
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 CTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCAT 1080

Qy        1084 GCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGAC 1143
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 GCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGAC 1140

Qy        1144 CTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGC 1203
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 CTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGC 1200

Qy        1204 TGAGCCGGGCACCGAGGCCCGGAGACACTATGATGA------------------------ 1239
               ||||||||||||||||||||||||||||||||||||
Db        1201 TGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGTAAGGCCTTGGCAGCCAGCAG 1260

Qy        1240 ------------------------------------------------------------ 1239

Db        1261 AGGCTGGTGTGGGAGCCGCCCACCAGAGACGACACTCGGGGCTGTGTCTGGGCTGGTGCC 1320
```

```
Qy       1240 ------------------------------------------------------------ 1239

Db       1321 TCTCCATCCTGGCCCCGACTTCTCTGTCAGGAAAGTGGGGATGGACCCCATCTGCATACA 1380

Qy       1240 ------------------------------------------------------------ 1239

Db       1381 CGGCTTCTCATGGGTGTGGAACATCTCTGCTTGCGGTTTCAGGAAGGCCTCTGGCTGCTC 1440

Qy       1240 ------------------------------------------------------------ 1239

Db       1441 TAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGACAGAAGGAAAGGCGGAGCTTATTCAA 1500

Qy       1240 ------------------------------------------------------------ 1239

Db       1501 AGTCTAGAGGGAGTGGAGGAGTTAAGGCTGGATTTCAGATCTGCCTGGTTCCAGCCGCAG 1560

Qy       1240 ------------------------------------------------------------ 1239

Db       1561 TGTGCCCTCTGCTCCCCCAACGACTTTCCAAATAATCTCACCAGCGCCTTCCAGCTCAGG 1620

Qy       1240 ------------------------------------------------------------ 1239

Db       1621 CGTCCTAGAAGCGTCTTGAAGCCTATGGCCAGCTGTCTTTGTGTTCCCTCTCACCCGCCT 1680

Qy       1240 ------------------------------------------------------------ 1239

Db       1681 GTCCTCACAGCTGAGACTCCCAGGAAACCTTCAGACTACCTTCCTCTGCCTTCAGCAAGG 1740

Qy       1240 ------------------------------------------------------------ 1239

Db       1741 GGCGTTGCCCACATTCTCTGAGGGTCAGTGGAAGAACCTAGACTCCCATTGCTAGAGGTA 1800

Qy       1240 ------------------------------------------------------------ 1239

Db       1801 GAAAGGGGAAGGGTGCTGGGGAGCAGGGCTGGTCCACAGCAGGTCTCGTGCAGCAGGTAC 1860

Qy       1240 ------------------------------------------------------------ 1239

Db       1861 CTGTGGTTCCGCCTTCTCATCTCCCTGAGACTGCTCCGACCCTTCCCTCCCAGGCTCTGT 1920

Qy       1240 ----------------------AGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGT 1277
                                    ||||||||||||||||||||||||||||||||||||||
Db       1921 CTGATGGCCCCTCTCCCTCTGCAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGT 1980

Qy       1278 GCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTC 1337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTC 2040

Qy       1338 GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT 1397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT 2100

Qy       1398 CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCC 1457
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCC 2160

Qy       1458 TGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGC 1517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 TGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGC 2220

Qy       1518 CCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCC 1577
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCC 2280

Qy       1578 TGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTG 1637
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 TGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTG 2340

Qy       1638 GCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTG 1697
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTG 2400

Qy       1698 TGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACC 1757
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACC 2460

Qy       1758 TCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCT 1817
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 TCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCT 2520

Qy       1818 CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTC 1877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTC 2580

Qy       1878 TGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT 1937
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 TGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT 2640

Qy       1938 CAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCC 1997
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCC 2700

Qy       1998 CCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAG 2057
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 CCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAG 2760

Qy       2058 TAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2117
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2820

Qy       2118 TGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAG 2177
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAG 2880

Qy       2178 GGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGG 2237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGG 2940

Qy       2238 GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACA 2297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACA 3000

Qy       2298 CCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGC 2357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGC 3060

Qy       2358 CCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACAC 2417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACAC 3120

Qy       2418 TCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAAC 2475
              ||||||||||||||||||||||||  | ||||||||||||||||||||||||||||||||
Db       3121 TCCTCCATGGGATTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAAC 3180

Qy       2476 ACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCT 2535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 ACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCT 3240

Qy       2536 TACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCAT 2595
              |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db       3241 TACCTTTTATCAGGATGT-GCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCAT 3299

Qy       2596 TTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGT 2655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3300 TTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGT 3359

Qy       2656 GTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAG 2715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3360 GTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAG 3419

Qy       2716 CTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAAT 2775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3420 CTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAAT 3479

Qy       2776 GCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTAC 2835
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3480 GCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTAC 3539

Qy          2836 CCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT 2895
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3540 CCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT 3599

Qy          2896 AACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCACTTCCACTCCCCTCTACTCTC 2955
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3600 AACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCACTTCCACTCCCCTCTACTCTC 3659

Qy          2956 TCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTA 3015
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3660 TCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTA 3719

Qy          3016 CCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCA 3075
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3720 CCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCA 3779

Qy          3076 CAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGG 3135
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3780 CAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGG 3839

Qy          3136 GAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTC 3195
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3840 GAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTC 3899

Qy          3196 AGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATAT 3255
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3900 AGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATAT 3959

Qy          3256 TTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTC 3315
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3960 TTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTC 4019

Qy          3316 TTATATGTTTAAAAA 3330
                 |||||||||||||||
Db          4020 TTATATGTTTAAAAA 4034
```

RESULT 19
US-10-012-896-703
; Sequence 703, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Wantanabe, Yoshihiro
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012,896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 703
; LENGTH: 2904
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-896-703

```
  Query Match             64.4%;  Score 2196.4;  DB 9;  Length 2904;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches 2560;  Conservative    0;  Mismatches    1;  Indels  343;  Gaps    2;

Qy           770 GTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC 829
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             1 GTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC 60

Qy           830 TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTG 889
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            61 TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTG 120

Qy           890 CTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCG 949
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           121 CTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCG 180

Qy           950 CTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGT 1009
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           181 CTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGT 240

Qy          1010 CCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAG 1069
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           241 CCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAG 300

Qy          1070 CTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGG 1129
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           301 CTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGG 360

Qy          1130 ATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAG 1189
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           361 ATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAG 420

Qy          1190 GGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAA--------- 1240
                 |||||||||||||||||||||||||||||||||||||||||||||||||||
Db           421 GGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGAAGGCCT 480

Qy          1241 ------------------------------------------------------------ 1240

Db           481 CTGGCTGCTCTAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGACAGAAGGAAAGGCGGA 540

Qy          1241 ------------------------------------------------------------ 1240

Db           541 GCTTATTCAAAGTCTAGAGGGAGTGGAGGAGTTAAGGCTGGATTTCAGATCTGCCTGGTT 600

Qy          1241 ------------------------------------------------------------ 1240

Db           601 CCAGCCGCAGTGTGCCCTCTGCTCCCCCAACGACTTTCCAAATAATCTCACCAGCGCCTT 660

Qy          1241 ------------------------------------------------------------ 1240

Db           661 CCAGCTCAGGCGTCCTAGAAGCGTCTTGAAGCCTATGGCCAGCTGTCTTTGTGTTCCCTC 720

Qy          1241 ------------------------------------------------------------ 1240

Db           721 TCACCCGCCTGTCCTCACAGCTGAGACTCCCAGGAAACCTTCAGACTACCTTCCTCTGCC 780

Qy          1241 --------------------------------GGCGTTCGGATGGCAGCCTGGGGCTGT 1268
                                                 ||||||||||||||||||||||||||||
Db           781 TTCAGCAAGGGGCGTTGCCCACATTCTCTGAGGGCGTTCGGATGGCAGCCTGGGGCTGT 840

Qy          1269 TCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGAT 1328
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           841 TCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGAT 900
```

```
Qy       1329 TCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCA 1388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCA 960

Qy       1389 CATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCT 1448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCT 1020

Qy       1449 TCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGG 1508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGG 1080

Qy       1509 TGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGA 1568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGA 1140

Qy       1569 CCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTG 1628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTG 1200

Qy       1629 GAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCT 1688
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCT 1260

Qy       1689 CCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCT 1748
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCT 1320

Qy       1749 GCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGT 1808
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGT 1380

Qy       1809 TTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAG 1868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAG 1440

Qy       1869 GCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGG 1928
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGG 1500

Qy       1929 CCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGT 1988
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGT 1560

Qy       1989 CCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTG 2048
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTG 1620

Qy       2049 CTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACA 2108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACA 1680

Qy       2109 GCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGG 2168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGG 1740

Qy       2169 CCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTG 2228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTG 1800

Qy       2229 GAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGT 2288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGT 1860

Qy       2289 TGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCA 2348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCA 1920

Qy       2349 TCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGG 2408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGG 1980

Qy       2409 ATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTG 2466
              ||||||||||||||||||||||||||||||||   | |||||||||||||||||||||||
Db       1981 ATGAAACACTCCTCCATGGGATTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTG 2040

Qy       2467 AGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCA 2526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 AGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCA 2100

Qy       2527 CCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGA 2586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGA 2160

Qy       2587 CACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAG 2646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 CACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAG 2220

Qy       2647 CTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCC 2706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCC 2280

Qy       2707 CTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCC 2766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCC 2340

Qy       2767 CCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAA 2826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 CCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAA 2400

Qy       2827 TGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAAC 2886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAAC 2460

Qy       2887 GGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCC 2946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCC 2520

Qy       2947 TCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAAC 3006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAAC 2580

Qy       3007 TTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGA 3066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 TTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGA 2640

Qy       3067 CCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCT 3126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCT 2700

Qy       3127 GTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTC 3186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTC 2760

Qy       3187 TTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGG 3246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGG 2820

Qy       3247 GGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTA 3306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 GGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTA 2880

Qy       3307 AAGGCTTTCTTATATGTTTAAAAA 3330
              ||||||||||||||||||||||||
Db       2881 AAGGCTTTCTTATATGTTTAAAAA 2904

RESULT 20
US-09-895-793-703
; Sequence 703, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
```

```
; APPLICANT:  Jiang, Yuqiu
; APPLICANT:  Kalos, Michael D.
; APPLICANT:  Retter, Marc W.
; APPLICANT:  Stolk, John A.
; APPLICANT:  Day, Craig H.
; APPLICANT:  Vedvick, Thomas S.
; APPLICANT:  Carter, Darrick
; APPLICANT:  Li, Samuel X.
; APPLICANT:  Wang, Aijun
; APPLICANT:  Skeiky, Yasir A.W.
; APPLICANT:  Hepler, William T.
; APPLICANT:  Henderson, Robert A.
; APPLICANT:  Hural, John
; APPLICANT:  McNeill, Patricia D.
; APPLICANT:  Houghton, Raymond L.
; APPLICANT:  Vinals de Bassols, Carlota
; APPLICANT:  Foy, Teresa
; APPLICANT:  Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE:  2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 703
;   LENGTH: 2904
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-895-793-703

  Query Match             64.4%;  Score 2196.4;  DB 9;  Length 2904;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches 2560;  Conservative    0;  Mismatches    1;  Indels  343;  Gaps    2;

Qy        770 GTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC 829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC 60

Qy        830 TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTG 889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTG 120

Qy        890 CTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCG 949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCG 180

Qy        950 CTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGT 1009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGT 240

Qy       1010 CCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAG 1069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAG 300

Qy       1070 CTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGG 1129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGG 360

Qy       1130 ATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAG 1189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAG 420

Qy       1190 GGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAA--------- 1240
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGAAGGCCT 480

Qy       1241 ------------------------------------------------------------ 1240

Db        481 CTGGCTGCTCTAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGACAGAAGGAAAGGCGGA 540

Qy       1241 ------------------------------------------------------------ 1240

Db        541 GCTTATTCAAAGTCTAGAGGGAGTGGAGGAGTTAAGGCTGGATTTCAGATCTGCCTGGTT 600

Qy       1241 ------------------------------------------------------------ 1240

Db        601 CCAGCCGCAGTGTGCCCTCTGCTCCCCCAACGACTTTCCAAATAATCTCACCAGCGCCTT 660

Qy       1241 ------------------------------------------------------------ 1240

Db        661 CCAGCTCAGGCGTCCTAGAAGCGTCTTGAAGCCTATGGCCAGCTGTCTTTGTGTTCCCTC 720

Qy       1241 ------------------------------------------------------------ 1240

Db        721 TCACCCGCCTGTCCTCACAGCTGAGACTCCCAGGAAACCTTCAGACTACCTTCCTCTGCC 780

Qy       1241 --------------------------------GGCGTTCGGATGGCAGCCTGGGGCTGT 1268
                                              ||||||||||||||||||||||||||||
Db        781 TTCAGCAAGGGGCGTTGCCCACATTCTCTGAGGGCGTTCGGATGGCAGCCTGGGGCTGT 840

Qy       1269 TCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGAT 1328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGAT 900

Qy       1329 TCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCA 1388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCA 960

Qy       1389 CATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCT 1448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCT 1020

Qy       1449 TCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGG 1508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGG 1080

Qy       1509 TGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGA 1568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGA 1140

Qy       1569 CCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTG 1628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTG 1200

Qy       1629 GAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCT 1688
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCT 1260

Qy       1689 CCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCT 1748
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCT 1320

Qy       1749 GCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGT 1808
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGT 1380

Qy       1809 TTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAG 1868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAG 1440

Qy       1869 GCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGG 1928
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGG 1500

Qy       1929 CCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGT 1988
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGT 1560

Qy       1989 CCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTG 2048
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTG 1620

Qy       2049 CTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACA 2108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACA 1680
```

```
Qy       2109 GCTGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGG 2168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GCTGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGG 1740

Qy       2169 CCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTG 2228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTG 1800

Qy       2229 GAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGT 2288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGT 1860

Qy       2289 TGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCA 2348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCA 1920

Qy       2349 TCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGG 2408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGG 1980

Qy       2409 ATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTG 2466
              |||||||||||||||||||||||||||||||||  | |||||||||||||||||||||||
Db       1981 ATGAAACACTCCTCCATGGGATTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTG 2040

Qy       2467 AGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCA 2526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 AGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCA 2100

Qy       2527 CCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGA 2586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGA 2160

Qy       2587 CACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAG 2646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 CACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAG 2220

Qy       2647 CTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCC 2706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCC 2280

Qy       2707 CTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCC 2766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCC 2340

Qy       2767 CCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAA 2826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 CCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAA 2400

Qy       2827 TGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAAC 2886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAAC 2460

Qy       2887 GGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCC 2946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCC 2520

Qy       2947 TCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAAC 3006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAAC 2580

Qy       3007 TTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGA 3066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 TTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGA 2640

Qy       3067 CCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCT 3126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCT 2700

Qy       3127 GTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTC 3186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTC 2760

Qy       3187 TTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGG 3246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGG 2820

Qy       3247 GGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTA 3306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 GGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTA 2880

Qy       3307 AAGGCTTTCTTATATGTTTAAAAA 3330
              ||||||||||||||||||||||||
Db       2881 AAGGCTTTCTTATATGTTTAAAAA 2904
```

```
RESULT 21
US-09-895-814-703
; Sequence 703, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C26
;  CURRENT APPLICATION NUMBER: US/09/895,814
;  CURRENT FILING DATE:  2001-06-29
;  NUMBER OF SEQ ID NOS: 990
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 703
;   LENGTH: 2904
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-895-814-703

  Query Match             64.4%;  Score 2196.4;  DB 9;  Length 2904;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches 2560;  Conservative    0;  Mismatches    1;  Indels  343;  Gaps    2;

Qy        770 GTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC 829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC 60

Qy        830 TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTG 889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTG 120

Qy        890 CTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCG 949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCG 180

Qy        950 CTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGT 1009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGT 240
```

```
Qy       1010 CCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAG 1069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAG 300

Qy       1070 CTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGG 1129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGG 360

Qy       1130 ATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAG 1189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAG 420

Qy       1190 GGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAA--------- 1240
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGAAGGCCT 480

Qy       1241 ------------------------------------------------------------ 1240

Db        481 CTGGCTGCTCTAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGACAGAAGGAAAGGCGGA 540

Qy       1241 ------------------------------------------------------------ 1240

Db        541 GCTTATTCAAAGTCTAGAGGGAGTGGAGGAGTTAAGGCTGGATTTCAGATCTGCCTGGTT 600

Qy       1241 ------------------------------------------------------------ 1240

Db        601 CCAGCCGCAGTGTGCCCTCTGCTCCCCCAACGACTTTCCAAATAATCTCACCAGCGCCTT 660

Qy       1241 ------------------------------------------------------------ 1240

Db        661 CCAGCTCAGGCGTCCTAGAAGCGTCTTGAAGCCTATGGCCAGCTGTCTTTGTGTTCCCTC 720

Qy       1241 ------------------------------------------------------------ 1240

Db        721 TCACCCGCCTGTCCTCACAGCTGAGACTCCCAGGAAACCTTCAGACTACCTTCCTCTGCC 780

Qy       1241 ---------------------------------GGCGTTCGGATGGCAGCCTGGGGCTGT 1268
                                               |||||||||||||||||||||||||||
Db        781 TTCAGCAAGGGGCGTTGCCCACATTCTCTGAGGGCGTTCGGATGGCAGCCTGGGGCTGT 840

Qy       1269 TCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGAT 1328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGAT 900

Qy       1329 TCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCA 1388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCA 960

Qy       1389 CATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCT 1448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCT 1020

Qy       1449 TCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGG 1508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGG 1080

Qy       1509 TGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGA 1568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGA 1140

Qy       1569 CCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTG 1628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTG 1200

Qy       1629 GAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCT 1688
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCT 1260

Qy       1689 CCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCT 1748
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCT 1320

Qy       1749 GCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGT 1808
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGT 1380

Qy       1809 TTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAG 1868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAG 1440

Qy       1869 GCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGG 1928
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGG 1500

Qy       1929 CCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGT 1988
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGT 1560

Qy       1989 CCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTG 2048
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTG 1620

Qy       2049 CTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACA 2108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACA 1680

Qy       2109 GCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGG 2168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGG 1740

Qy       2169 CCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTG 2228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTG 1800

Qy       2229 GAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGT 2288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGT 1860

Qy       2289 TGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCA 2348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCA 1920

Qy       2349 TCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGG 2408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGG 1980

Qy       2409 ATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTG 2466
              |||||||||||||||||||||||||||||||||  | |||||||||||||||||||||||
Db       1981 ATGAAACACTCCTCCATGGGATTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTG 2040

Qy       2467 AGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCA 2526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 AGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCA 2100

Qy       2527 CCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGA 2586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGA 2160

Qy       2587 CACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAG 2646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 CACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAG 2220

Qy       2647 CTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCC 2706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCC 2280

Qy       2707 CTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCC 2766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCC 2340

Qy       2767 CCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAA 2826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 CCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAA 2400

Qy       2827 TGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAAC 2886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2401 TGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAAC 2460

Qy       2887 GGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCC 2946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCC 2520

Qy       2947 TCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAAC 3006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAAC 2580

Qy       3007 TTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGA 3066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 TTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGA 2640

Qy       3067 CCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCT 3126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCT 2700

Qy       3127 GTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTC 3186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTC 2760

Qy       3187 TTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGG 3246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGG 2820

Qy       3247 GGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTA 3306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 GGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTA 2880

Qy       3307 AAGGCTTTCTTATATGTTTAAAAA 3330
              ||||||||||||||||||||||||
Db       2881 AAGGCTTTCTTATATGTTTAAAAA 2904


RESULT 22
US-09-759-143-703
; Sequence 703, Application US/09759143
; Patent No. US20020022248A1
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C23
;  CURRENT APPLICATION NUMBER: US/09/759,143
;  CURRENT FILING DATE:  2001-01-12
;  NUMBER OF SEQ ID NOS:  934
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 703
;   LENGTH: 2904
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-759-143-703

  Query Match             64.4%;  Score 2196.4;  DB 10;  Length 2904;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches 2560;  Conservative    0;  Mismatches    1;  Indels  343;  Gaps    2;

Qy        770 GTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC 829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC 60

Qy        830 TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTG 889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTG 120

Qy        890 CTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCG 949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCG 180

Qy        950 CTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGT 1009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGT 240

Qy       1010 CCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAG 1069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAG 300

Qy       1070 CTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGG 1129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGG 360

Qy       1130 ATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAG 1189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAG 420

Qy       1190 GGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAA--------- 1240
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGAAGGCCT 480

Qy       1241 ------------------------------------------------------------ 1240

Db        481 CTGGCTGCTCTAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGACAGAAGGAAAGGCGGA 540

Qy       1241 ------------------------------------------------------------ 1240

Db        541 GCTTATTCAAAGTCTAGAGGGAGTGGAGGAGTTAAGGCTGGATTTCAGATCTGCCTGGTT 600

Qy       1241 ------------------------------------------------------------ 1240

Db        601 CCAGCCGCAGTGTGCCCTCTGCTCCCCCAACGACTTTCCAAATAATCTCACCAGCGCCTT 660

Qy       1241 ------------------------------------------------------------ 1240

Db        661 CCAGCTCAGGCGTCCTAGAAGCGTCTTGAAGCCTATGGCCAGCTGTCTTTGTGTTCCCTC 720

Qy       1241 ------------------------------------------------------------ 1240

Db        721 TCACCCGCCTGTCCTCACAGCTGAGACTCCCAGGAAACCTTCAGACTACCTTCCTCTGCC 780

Qy       1241 --------------------------------GGCGTTCGGATGGGCAGCCTGGGGCTGT 1268
                                              ||||||||||||||||||||||||||||
Db        781 TTCAGCAAGGGGCGTTGCCCACATTCTCTGAGGGCGTTCGGATGGGCAGCCTGGGGCTGT 840

Qy       1269 TCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGAT 1328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGAT 900

Qy       1329 TCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCA 1388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCA 960

Qy       1389 CATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCT 1448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCT 1020

Qy       1449 TCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGG 1508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGG 1080
```

```
Qy       1509 TGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGA 1568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGA 1140

Qy       1569 CCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTG 1628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTG 1200

Qy       1629 GAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCT 1688
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCT 1260

Qy       1689 CCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCT 1748
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCT 1320

Qy       1749 GCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGT 1808
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGT 1380

Qy       1809 TTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAG 1868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAG 1440

Qy       1869 GCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGG 1928
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGG 1500

Qy       1929 CCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGT 1988
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGT 1560

Qy       1989 CCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTG 2048
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTG 1620

Qy       2049 CTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACA 2108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACA 1680

Qy       2109 GCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGG 2168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGG 1740

Qy       2169 CCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTG 2228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTG 1800

Qy       2229 GAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGT 2288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGT 1860

Qy       2289 TGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCA 2348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCA 1920

Qy       2349 TCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGG 2408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGG 1980

Qy       2409 ATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTG 2466
              |||||||||||||||||||||||||||||||||  | |||||||||||||||||||||||
Db       1981 ATGAAACACTCCTCCATGGGATTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTG 2040

Qy       2467 AGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCA 2526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 AGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCA 2100

Qy       2527 CCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGA 2586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGA 2160

Qy       2587 CACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAG 2646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 CACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAG 2220

Qy       2647 CTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCC 2706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCC 2280

Qy       2707 CTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCC 2766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCC 2340

Qy       2767 CCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAA 2826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 CCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAA 2400

Qy       2827 TGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAAC 2886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAAC 2460

Qy       2887 GGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCC 2946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCC 2520

Qy       2947 TCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAAC 3006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAAC 2580

Qy       3007 TTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGA 3066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 TTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGA 2640

Qy       3067 CCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCT 3126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCT 2700

Qy       3127 GTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTC 3186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTC 2760

Qy       3187 TTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGG 3246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGG 2820

Qy       3247 GGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTA 3306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 GGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTA 2880

Qy       3307 AAGGCTTTCTTATATGTTTAAAAA 3330
              ||||||||||||||||||||||||
Db       2881 AAGGCTTTCTTATATGTTTAAAAA 2904
```

```
RESULT 23
US-09-780-669-703
; Sequence 703, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
```

```
;  APPLICANT:  Hepler, William
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C24
;  CURRENT APPLICATION NUMBER: US/09/780,669
;  CURRENT FILING DATE:  2001-02-09
;  NUMBER OF SEQ ID NOS:  943
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 703
;   LENGTH: 2904
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-780-669-703

  Query Match             64.4%;  Score 2196.4;  DB 10;  Length 2904;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches 2560;  Conservative    0;  Mismatches    1;  Indels  343;  Gaps    2;

Qy        770 GTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC 829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC 60

Qy        830 TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTG 889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTG 120

Qy        890 CTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCG 949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCG 180

Qy        950 CTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGT 1009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGT 240

Qy       1010 CCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAG 1069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAG 300

Qy       1070 CTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGG 1129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGG 360

Qy       1130 ATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAG 1189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAG 420

Qy       1190 GGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAA--------- 1240
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGAAGGCCT 480

Qy       1241 ------------------------------------------------------------ 1240

Db        481 CTGGCTGCTCTAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGACAGAAGGAAAGGCGGA 540

Qy       1241 ------------------------------------------------------------ 1240

Db        541 GCTTATTCAAAGTCTAGAGGGAGTGGAGGAGTTAAGGCTGGATTTCAGATCTGCCTGGTT 600

Qy       1241 ------------------------------------------------------------ 1240

Db        601 CCAGCCGCAGTGTGCCCTCTGCTCCCCCAACGACTTTCCAAATAATCTCACCAGCGCCTT 660

Qy       1241 ------------------------------------------------------------ 1240

Db        661 CCAGCTCAGGCGTCCTAGAAGCGTCTTGAAGCCTATGGCCAGCTGTCTTTGTGTTCCCTC 720

Qy       1241 ------------------------------------------------------------ 1240

Db        721 TCACCCGCCTGTCCTCACAGCTGAGACTCCCAGGAAACCTTCAGACTACCTTCCTCTGCC 780

Qy       1241 --------------------------------GGCGTTCGGATGGCAGCCTGGGGCTGT 1268
                                              ||||||||||||||||||||||||||||
Db        781 TTCAGCAAGGGGCGTTGCCCACATTCTCTGAGGGCGTTCGGATGGCAGCCTGGGGCTGT 840

Qy       1269 TCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGAT 1328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGAT 900

Qy       1329 TCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCA 1388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCA 960

Qy       1389 CATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCT 1448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCT 1020

Qy       1449 TCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGG 1508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGG 1080

Qy       1509 TGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGA 1568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGA 1140

Qy       1569 CCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTG 1628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTG 1200

Qy       1629 GAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCT 1688
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCT 1260

Qy       1689 CCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCT 1748
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCT 1320

Qy       1749 GCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGT 1808
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGT 1380

Qy       1809 TTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAG 1868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAG 1440

Qy       1869 GCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGG 1928
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGG 1500

Qy       1929 CCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGT 1988
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGT 1560

Qy       1989 CCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTG 2048
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTG 1620

Qy       2049 CTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACA 2108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACA 1680

Qy       2109 GCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGG 2168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGG 1740

Qy       2169 CCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTG 2228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTG 1800

Qy       2229 GAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGT 2288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGT 1860

Qy       2289 TGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCA 2348
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGAGACACACCTAGAGAAGGGTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCA 1920

Qy       2349 TCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGG 2408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGG 1980

Qy       2409 ATGAAACACTCCTCCATGGATTTGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTG 2466
              ||||||||||||||||||||||||||||||||  | ||||||||||||||||||||||||
Db       1981 ATGAAACACTCCTCCATGGATTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTG 2040

Qy       2467 AGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCA 2526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 AGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCA 2100

Qy       2527 CCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGA 2586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGA 2160

Qy       2587 CACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAG 2646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 CACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAG 2220

Qy       2647 CTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCC 2706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCC 2280

Qy       2707 CTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCC 2766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCC 2340

Qy       2767 CCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAA 2826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 CCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAA 2400

Qy       2827 TGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAAC 2886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAAC 2460

Qy       2887 GGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCC 2946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCC 2520

Qy       2947 TCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAAC 3006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAAC 2580

Qy       3007 TTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGA 3066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 TTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGA 2640

Qy       3067 CCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCT 3126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCT 2700

Qy       3127 GTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTC 3186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTC 2760

Qy       3187 TTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGG 3246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGG 2820

Qy       3247 GGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTA 3306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 GGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTA 2880

Qy       3307 AAGGCTTTCTTATATGTTTAAAAA 3330
              ||||||||||||||||||||||||
Db       2881 AAGGCTTTCTTATATGTTTAAAAA 2904
```

RESULT 24

```
US-09-822-827-703
; Sequence 703, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.534C1
;  CURRENT APPLICATION NUMBER: US/09/822,827
;  CURRENT FILING DATE:  2001-03-28
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 703
;   LENGTH: 2904
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-822-827-703

  Query Match             64.4%;  Score 2196.4;  DB 10;  Length 2904;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches 2560;  Conservative    0;  Mismatches    1;  Indels  343;  Gaps    2;

Qy        770 GTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC 829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC 60

Qy        830 TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTG 889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTG 120

Qy        890 CTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCG 949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCG 180

Qy        950 CTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGT 1009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGT 240

Qy       1010 CCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAG 1069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAG 300

Qy       1070 CTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGG 1129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGG 360

Qy       1130 ATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAG 1189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAG 420

Qy       1190 GGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAA--------- 1240
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGAAGGCCT 480

Qy       1241 ------------------------------------------------------------ 1240

Db        481 CTGGCTGCTCTAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGACAGAAGGAAAGGCGGA 540

Qy       1241 ------------------------------------------------------------ 1240

Db        541 GCTTATTCAAAGTCTAGAGGGAGTGGAGGAGTTAAGGCTGGATTTCAGATCTGCCTGGTT 600

Qy       1241 ------------------------------------------------------------ 1240

Db        601 CCAGCCGCAGTGTGCCCTCTGCTCCCCCAACGACTTTCCAAATAATCTCACCAGCGCCTT 660

Qy       1241 ------------------------------------------------------------ 1240

Db        661 CCAGCTCAGGCGTCCTAGAAGCGTCTTGAAGCCTATGGCCAGCTGTCTTTGTGTTCCCTC 720

Qy       1241 ------------------------------------------------------------ 1240

Db        721 TCACCCGCCTGTCCTCACAGCTGAGACTCCCAGGAAACCTTCAGACTACCTTCCTCTGCC 780
```

```
Qy   1241 --------------------------------GGCGTTCGGATGGGCAGCCTGGGGCTGT 1268
                                          ||||||||||||||||||||||||||||
Db    781 TTCAGCAAGGGGCGTTGCCCACATTCTCTGAGGGCGTTCGGATGGGCAGCCTGGGGCTGT 840

Qy   1269 TCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGAT 1328
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGAT 900

Qy   1329 TCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCA 1388
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCA 960

Qy   1389 CATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCT 1448
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCT 1020

Qy   1449 TCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGG 1508
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGG 1080

Qy   1509 TGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGA 1568
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGA 1140

Qy   1569 CCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTG 1628
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTG 1200

Qy   1629 GAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGCCTCTGCCTGTGATGTCT 1688
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGCCTCTGCCTGTGATGTCT 1260

Qy   1689 CCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCT 1748
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCT 1320

Qy   1749 GCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGT 1808
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGT 1380

Qy   1809 TTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAG 1868
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAG 1440

Qy   1869 GCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGG 1928
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGG 1500

Qy   1929 CCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGT 1988
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGT 1560

Qy   1989 CCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTG 2048
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTG 1620

Qy   2049 CTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACA 2108
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 CTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACA 1680

Qy   2109 GCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGG 2168
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGG 1740

Qy   2169 CCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTG 2228
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 CCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTG 1800

Qy   2229 GAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGT 2288
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 GAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGT 1860

Qy   2289 TGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCA 2348
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCA 1920

Qy   2349 TCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGG 2408
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 TCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGG 1980

Qy   2409 ATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTG 2466
          |||||||||||||||||||||||||||||||||  | |||||||||||||||||||||||
Db   1981 ATGAAACACTCCTCCATGGGATTTGAACATATGAAGTTATTTGTAGGGGAAGAGTCCTG 2040

Qy   2467 AGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCA 2526
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 AGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCA 2100

Qy   2527 CCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGA 2586
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 CCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGA 2160

Qy   2587 CACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAG 2646
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 CACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAG 2220

Qy   2647 CTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCC 2706
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 CTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCC 2280

Qy   2707 CTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCC 2766
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCC 2340

Qy   2767 CCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAA 2826
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 CCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAA 2400

Qy   2827 TGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAAC 2886
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAAC 2460

Qy   2887 GGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCC 2946
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 GGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCC 2520

Qy   2947 TCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAAC 3006
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 TCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAAC 2580

Qy   3007 TTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGA 3066
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 TTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGA 2640

Qy   3067 CCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCT 3126
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 CCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCT 2700

Qy   3127 GTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTC 3186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 GTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTC 2760

Qy   3187 TTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGG 3246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 TTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGG 2820

Qy   3247 GGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTA 3306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 GGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTA 2880

Qy   3307 AAGGCTTTCTTATATGTTTAAAAA 3330
          ||||||||||||||||||||||||
Db   2881 AAGGCTTTCTTATATGTTTAAAAA 2904
```

```
RESULT 25
US-10-012-896-702
; Sequence 702, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Wantanabe, Yoshihiro
;  APPLICANT:  Meagher, Madeleine Joy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C27
;  CURRENT APPLICATION NUMBER: US/10/012,896
;  CURRENT FILING DATE:  2001-12-10
;  NUMBER OF SEQ ID NOS: 1011
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 702
;   LENGTH: 4894
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-012-896-702

  Query Match             62.8%;  Score 2142.8;  DB 9;  Length 4894;
  Best Local Similarity   80.1%;  Pred. No. 0;
  Matches 2872;  Conservative    1;  Mismatches    3;  Indels  708;  Gaps    2;

Qy        455 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1311 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 1370

Qy        515 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1371 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 1430

Qy        575 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1431 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 1490

Qy        635 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 694
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1491 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 1550

Qy        695 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1551 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 1610

Qy        755 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 814
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1611 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 1670

Qy        815 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 874
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1671 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 1730

Qy        875 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 934
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1731 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 1790

Qy        935 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 994
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1791 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 1850

Qy        995 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1054
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1851 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1910

Qy       1055 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1911 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1970

Qy       1115 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 1174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1971 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 2030

Qy       1175 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 1234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2031 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 2090

Qy       1235 GATGA------------------------------------------------------- 1239
              |||||
Db       2091 GATGAAGGTAAGGCCTTGGCAGCCAGCAGAGGCTGGTGTGGGAGCCGCCCACCAGAGACG 2150

Qy       1240 ------------------------------------------------------------ 1239

Db       2151 ACACTCGGGGCTGTGTCTGGGCTGGTGCCTCTCCATCCTGGCCCCGACTTCTCTGTCAGG 2210

Qy       1240 ------------------------------------------------------------ 1239

Db       2211 AAAGTGGGGATGGACCCCATCTGCATACACGGCTTCTCATGGGTGTGGAACATCTCTGCT 2270

Qy       1240 ------------------------------------------------------------ 1239

Db       2271 TGCGGTTTCAGGAAGGCCTCTGGCTGCTCTAGGAGTCTGATCAGAGTCGTTGCCCCAGTT 2330

Qy       1240 ------------------------------------------------------------ 1239

Db       2331 TGACAGAAGGAAAGGCGGAGCTTATTCAAAGTCTAGAGGGAGTGGAGGAGTTAAGGCTGG 2390

Qy       1240 ------------------------------------------------------------ 1239

Db       2391 ATTTCAGATCTGCCTGGTTCCAGCCGCAGTGTGCCCTCTGCTCCCCCAACGACTTTCCAA 2450

Qy       1240 ------------------------------------------------------------ 1239

Db       2451 ATAATCTCACCAGCGCCTTCCAGCTCAGGCGTCCTAGAAGCGTCTTGAAGCCTATGGCCA 2510

Qy       1240 ------------------------------------------------------------ 1239

Db       2511 GCTGTCTTTGTGTTCCCTCTCACCCGCCTGTCCTCACAGCTGAGACTCCCAGGAAACCTT 2570

Qy       1240 ------------------------------------------------------------ 1239

Db       2571 CAGACTACCTTCCTCTGCCTTCAGCAAGGGGCGTTGCCCACATTCTCTGAGGGTCAGTGG 2630

Qy       1240 ------------------------------------------------------------ 1239

Db       2631 AAGAACCTAGACTCCCATTGCTAGAGGTAGAAAGGGGAAGGGTGCTGGGGAGCAGGGCTG 2690

Qy       1240 ------------------------------------------------------------ 1239

Db       2691 GTCCACAGCAGGTCTCGTGCAGCAGGTACCTGTGGTTCCGCCTTCTCATCTCCCTGAGAC 2750

Qy       1240 ---------------------------------------------------AGGCGTTCG 1248
                                                                 |||||||||
Db       2751 TGCTCCGACCCTTCCCTCCCAGGCTCTGTCTGATGGCCCCTCTCCCTCTGCAGGCGTTCG 2810
```

```
Qy       1249 GATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCAT 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2811 GATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCAT 2870

Qy       1309 GGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 1368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2871 GGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 2930

Qy       1369 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 1428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2931 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 2990

Qy       1429 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 1488
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2991 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 3050

Qy       1489 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 1548
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3051 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 3110

Qy       1549 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 1608
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3111 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 3170

Qy       1609 TAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 1668
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db       3171 TAATGGACACGTGGGTGYTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 3230

Qy       1669 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 1728
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3231 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 3290

Qy       1729 GGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTC 1788
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3291 GGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTC 3350

Qy       1789 CCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGC 1848
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3351 CCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGC 3410

Qy       1849 CTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT 1908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3411 CTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT 3470

Qy       1909 ATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG 1968
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3471 ATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG 3530

Qy       1969 GAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGG 2028
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db       3531 GAGGGCCTGCCTCACTGGGTCCCAGCTCCCTGCTCCTGTTAGCCCCATGGGGCTGCCGGG 3590

Qy       2029 CTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTG 2088
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3591 CTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTG 3650

Qy       2089 CTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTC 2148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3651 CTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTC 3710

Qy       2149 TAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGC 2208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3711 TAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGC 3770

Qy       2209 CAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG 2268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3771 CAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG 3830

Qy       2269 TTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAA 2328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3831 TTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAA 3890

Qy       2329 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCT 2388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3891 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCT 3950

Qy       2389 CTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTA 2446
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||  | |||
Db       3951 CTTGCATGGGAGTTTCTAGGATGAAACACTCCACCATGGGATTTGAACATATGAAAGTTA 4010

Qy       2447 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 2506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4011 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 4070

Qy       2507 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 2566
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4071 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 4130

Qy       2567 TTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGT 2626
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4131 TTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGT 4190

Qy       2627 AGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 2686
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4191 AGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 4250

Qy       2687 ATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCT 2746
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4251 ATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCT 4310

Qy       2747 TCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCA 2806
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4311 TCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCA 4370

Qy       2807 TCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGG 2866
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4371 TCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGG 4430

Qy       2867 GTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGT 2926
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4431 GTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGT 4490

Qy       2927 TCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 2986
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4491 TCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 4550

Qy       2987 AAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCC 3046
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4551 AAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCC 4610

Qy       3047 TGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATC 3106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4611 TGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATC 4670

Qy       3107 TCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCC 3166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4671 TCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCC 4730

Qy       3167 TGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATG 3226
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4731 TGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATG 4790

Qy       3227 TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAA 3286
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4791 TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAA 4850

Qy       3287 TGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAA 3330
              ||||||||||||||||||||||||||||||||||||||||||||
Db       4851 TGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAA 4894
```

RESULT 26
US-09-895-793-702
; Sequence 702, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:

```
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.534C2
;  CURRENT APPLICATION NUMBER: US/09/895,793
;  CURRENT FILING DATE:  2001-06-29
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 702
;   LENGTH: 4894
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-895-793-702

  Query Match             62.8%;  Score 2142.8;  DB 9;  Length 4894;
  Best Local Similarity   80.1%;  Pred. No. 0;
  Matches 2872;  Conservative    1;  Mismatches    3;  Indels  708;  Gaps    2;

Qy        455 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1311 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 1370

Qy        515 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1371 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 1430

Qy        575 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1431 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 1490

Qy        635 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 694
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1491 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 1550

Qy        695 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1551 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 1610

Qy        755 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 814
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1611 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 1670

Qy        815 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 874
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1671 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 1730

Qy        875 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 934
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1731 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 1790

Qy        935 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 994
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1791 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 1850

Qy        995 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1054
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1851 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1910

Qy       1055 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1911 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1970

Qy       1115 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 1174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1971 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 2030

Qy       1175 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 1234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2031 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 2090

Qy       1235 GATGA------------------------------------------------------- 1239
              |||||
Db       2091 GATGAAGGTAAGGCCTTGGCAGCCAGCAGAGGCTGGTGTGGGAGCCGCCCACCAGAGACG 2150

Qy       1240 ------------------------------------------------------------ 1239

Db       2151 ACACTCGGGGCTGTGTCTGGGCTGGTGCCTCTCCATCCTGGCCCCGACTTCTCTGTCAGG 2210

Qy       1240 ------------------------------------------------------------ 1239

Db       2211 AAAGTGGGGATGGACCCCATCTGCATACACGGCTTCTCATGGGTGTGGAACATCTCTGCT 2270

Qy       1240 ------------------------------------------------------------ 1239

Db       2271 TGCGGTTTCAGGAAGGCCTCTGGCTGCTCTAGGAGTCTGATCAGAGTCGTTGCCCCAGTT 2330

Qy       1240 ------------------------------------------------------------ 1239

Db       2331 TGACAGAAGGAAAGGCGGAGCTTATTCAAAGTCTAGAGGGAGTGGAGGAGTTAAGGCTGG 2390

Qy       1240 ------------------------------------------------------------ 1239

Db       2391 ATTTCAGATCTGCCTGGTTCCAGCCGCAGTGTGCCCTCTGCTCCCCCAACGACTTTCCAA 2450

Qy       1240 ------------------------------------------------------------ 1239

Db       2451 ATAATCTCACCAGCGCCTTCCAGCTCAGGCGTCCTAGAAGCGTCTTGAAGCCTATGGCCA 2510

Qy       1240 ------------------------------------------------------------ 1239

Db       2511 GCTGTCTTTGTGTTCCCTCTCACCCGCCTGTCCTCACAGCTGAGACTCCCAGGAAACCTT 2570

Qy       1240 ------------------------------------------------------------ 1239

Db       2571 CAGACTACCTTCCTCTGCCTTCAGCAAGGGGCGTTGCCCACATTCTCTGAGGGTCAGTGG 2630

Qy       1240 ------------------------------------------------------------ 1239

Db       2631 AAGAACCTAGACTCCCATTGCTAGAGGTAGAAAGGGGAAGGGTGCTGGGGAGCAGGGCTG 2690

Qy       1240 ------------------------------------------------------------ 1239

Db       2691 GTCCACAGCAGGTCTCGTGCAGCAGGTACCTGTGGTTCCGCCTTCTCATCTCCCTGAGAC 2750

Qy       1240 ---------------------------------------------------AGGCGTTCG 1248
                                                                 |||||||||
Db       2751 TGCTCCGACCCTTCCCTCCCAGGCTCTGTCTGATGGCCCCTCTCCCTCTGCAGGCGTTCG 2810

Qy       1249 GATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCAT 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2811 GATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCAT 2870

Qy       1309 GGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 1368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2871 GGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 2930
```

```
Qy   1369 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 1428
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2931 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 2990

Qy   1429 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 1488
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2991 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 3050

Qy   1489 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 1548
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3051 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 3110

Qy   1549 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 1608
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3111 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 3170

Qy   1609 TAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 1668
          |||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db   3171 TAATGGACACGTGGGTGYTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 3230

Qy   1669 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 1728
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3231 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 3290

Qy   1729 GGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTC 1788
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3291 GGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTC 3350

Qy   1789 CCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGC 1848
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3351 CCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGC 3410

Qy   1849 CTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT 1908
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3411 CTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT 3470

Qy   1909 ATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG 1968
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3471 ATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG 3530

Qy   1969 GAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGG 2028
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   3531 GAGGGCCTGCCTCACTGGGTCCCAGCTCCCTGCTCCTGTTAGCCCCATGGGGCTGCCGGG 3590

Qy   2029 CTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTG 2088
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3591 CTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTG 3650

Qy   2089 CTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTC 2148
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3651 CTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTC 3710

Qy   2149 TAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGGTTTCAGTCTGGACTTATACAGGGAGGC 2208
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3711 TAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGGTTTCAGTCTGGACTTATACAGGGAGGC 3770

Qy   2209 CAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG 2268
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3771 CAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG 3830

Qy   2269 TTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTGGGAGCTGAATAAA 2328
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3831 TTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTGGGAGCTGAATAAA 3890

Qy   2329 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCT 2388
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3891 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCT 3950

Qy   2389 CTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTA 2446
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||  | |||
Db   3951 CTTGCATGGGAGTTTCTAGGATGAAACACTCCACCATGGGATTTGAACATATGAAAGTTA 4010

Qy   2447 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 2506
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4011 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 4070

Qy   2507 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 2566
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4071 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 4130

Qy   2567 TTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGT 2626
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4131 TTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGT 4190

Qy   2627 AGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 2686
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4191 AGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 4250

Qy   2687 ATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCT 2746
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4251 ATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCT 4310

Qy   2747 TCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCA 2806
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4311 TCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCA 4370

Qy   2807 TCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGG 2866
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4371 TCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGG 4430

Qy   2867 GTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGT 2926
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4431 GTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGT 4490

Qy   2927 TCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 2986
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4491 TCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 4550

Qy   2987 AAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCC 3046
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4551 AAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCC 4610

Qy   3047 TGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATC 3106
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4611 TGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATC 4670

Qy   3107 TCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCC 3166
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4671 TCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCC 4730

Qy   3167 TGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATG 3226
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4731 TGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATG 4790

Qy   3227 TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAA 3286
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4791 TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAA 4850

Qy   3287 TGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAA 3330
          ||||||||||||||||||||||||||||||||||||||||||||
Db   4851 TGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAA 4894
```

```
RESULT 27
US-09-895-814-702
; Sequence 702, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
```

```
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C26
;  CURRENT APPLICATION NUMBER: US/09/895,814
;  CURRENT FILING DATE:  2001-06-29
;  NUMBER OF SEQ ID NOS: 990
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 702
;   LENGTH: 4894
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-895-814-702

  Query Match             62.8%;  Score 2142.8;  DB 9;  Length 4894;
  Best Local Similarity   80.1%;  Pred. No. 0;
  Matches 2872;  Conservative    1;  Mismatches    3;  Indels  708;  Gaps    2;

Qy        455 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1311 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 1370

Qy        515 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1371 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 1430

Qy        575 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1431 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 1490

Qy        635 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 694
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1491 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 1550

Qy        695 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1551 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 1610

Qy        755 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 814
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1611 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 1670

Qy        815 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 874
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1671 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 1730

Qy        875 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 934
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1731 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 1790

Qy        935 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 994
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1791 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 1850

Qy        995 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1054
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1851 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1910

Qy       1055 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1911 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1970

Qy       1115 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTACACGGATTTCGTGGGC 1174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1971 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTACACGGATTTCGTGGGC 2030

Qy       1175 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 1234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2031 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 2090

Qy       1235 GATGA------------------------------------------------------- 1239
              |||||
Db       2091 GATGAAGGTAAGGCCTTGGCAGCCAGCAGAGGCTGGTGTGGGAGCCGCCCACCAGAGACG 2150

Qy       1240 ------------------------------------------------------------ 1239

Db       2151 ACACTCGGGGCTGTGTCTGGGCTGGTGCCTCTCCATCCTGGCCCCGACTTCTCTGTCAGG 2210

Qy       1240 ------------------------------------------------------------ 1239

Db       2211 AAAGTGGGGATGGACCCCATCTGCATACACGGCTTCTCATGGGTGTGGAACATCTCTGCT 2270

Qy       1240 ------------------------------------------------------------ 1239

Db       2271 TGCGGTTTCAGGAAGGCCTCTGGCTGCTCTAGGAGTCTGATCAGAGTCGTTGCCCCAGTT 2330

Qy       1240 ------------------------------------------------------------ 1239

Db       2331 TGACAGAAGGAAAGGCGGAGCTTATTCAAAGTCTAGAGGGAGTGGAGGAGTTAAGGCTGG 2390

Qy       1240 ------------------------------------------------------------ 1239

Db       2391 ATTTCAGATCTGCCTGGTTCCAGCCGCAGTGTGCCCTCTGCTCCCCCAACGACTTTCCAA 2450

Qy       1240 ------------------------------------------------------------ 1239

Db       2451 ATAATCTCACCAGCGCCTTCCAGCTCAGGCGTCCTAGAAGCGTCTTGAAGCCTATGGCCA 2510

Qy       1240 ------------------------------------------------------------ 1239

Db       2511 GCTGTCTTTGTGTTCCCTCTCACCCGCCTGTCCTCACAGCTGAGACTCCCAGGAAACCTT 2570

Qy       1240 ------------------------------------------------------------ 1239

Db       2571 CAGACTACCTTCCTCTGCCTTCAGCAAGGGGCGTTGCCCACATTCTCTGAGGGTCAGTGG 2630

Qy       1240 ------------------------------------------------------------ 1239

Db       2631 AAGAACCTAGACTCCCATTGCTAGAGGTAGAAAGGGGAAGGGTGCTGGGGAGCAGGGCTG 2690

Qy       1240 ------------------------------------------------------------ 1239

Db       2691 GTCCACAGCAGGTCTCGTGCAGCAGGTACCTGTGGTTCCGCCTTCTCATCTCCCTGAGAC 2750

Qy       1240 ---------------------------------------------------AGGCGTTCG 1248
                                                                 |||||||||
Db       2751 TGCTCCGACCCTTCCCTCCCAGGCTCTGTCTGATGGCCCCTCTCCCTCTGCAGGCGTTCG 2810

Qy       1249 GATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCAT 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2811 GATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCAT 2870

Qy       1309 GGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 1368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2871 GGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 2930

Qy       1369 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 1428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2931 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 2990

Qy       1429 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 1488
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     2991 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 3050

Qy     1489 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 1548
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3051 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 3110

Qy     1549 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 1608
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3111 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 3170

Qy     1609 TAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 1668
            |||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db     3171 TAATGGACACGTGGGTGYTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 3230

Qy     1669 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 1728
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3231 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 3290

Qy     1729 GGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTC 1788
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3291 GGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTC 3350

Qy     1789 CCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGC 1848
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3351 CCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGC 3410

Qy     1849 CTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT 1908
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3411 CTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT 3470

Qy     1909 ATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG 1968
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3471 ATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG 3530

Qy     1969 GAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGG 2028
            |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db     3531 GAGGGCCTGCCTCACTGGGTCCCAGCTCCCTGCTCCTGTTAGCCCCATGGGGCTGCCGGG 3590

Qy     2029 CTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTG 2088
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3591 CTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTG 3650

Qy     2089 CTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTC 2148
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3651 CTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTC 3710

Qy     2149 TAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGC 2208
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3711 TAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGC 3770

Qy     2209 CAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG 2268
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3771 CAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG 3830

Qy     2269 TTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAA 2328
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3831 TTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAA 3890

Qy     2329 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCT 2388
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3891 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCT 3950

Qy     2389 CTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTA 2446
            |||||||||||||||||||||||||||||||| ||||||||||||||||||||  | |||
Db     3951 CTTGCATGGGAGTTTCTAGGATGAAACACTCCACCATGGGATTTGAACATATGAAAGTTA 4010

Qy     2447 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 2506
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4011 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 4070

Qy     2507 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 2566
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4071 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 4130

Qy     2567 TTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGT 2626
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4131 TTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGT 4190

Qy     2627 AGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 2686
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4191 AGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 4250

Qy     2687 ATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCT 2746
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4251 ATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCT 4310

Qy     2747 TCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCA 2806
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4311 TCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCA 4370

Qy     2807 TCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGG 2866
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4371 TCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGG 4430

Qy     2867 GTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGT 2926
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4431 GTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGT 4490

Qy     2927 TCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 2986
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4491 TCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 4550

Qy     2987 AAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCC 3046
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4551 AAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCC 4610

Qy     3047 TGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATC 3106
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4611 TGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATC 4670

Qy     3107 TCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCC 3166
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4671 TCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCC 4730

Qy     3167 TGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATG 3226
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4731 TGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATG 4790

Qy     3227 TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAA 3286
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4791 TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAA 4850

Qy     3287 TGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAA 3330
            ||||||||||||||||||||||||||||||||||||||||||||
Db     4851 TGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAA 4894
```

```
RESULT 28
US-09-759-143-702
; Sequence 702, Application US/09759143
; Patent No. US20020022248A1
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
```

```
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C23
;  CURRENT APPLICATION NUMBER: US/09/759,143
;  CURRENT FILING DATE:  2001-01-12
;  NUMBER OF SEQ ID NOS:  934
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 702
;   LENGTH: 4894
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-759-143-702

  Query Match             62.8%;  Score 2142.8;  DB 10;  Length 4894;
  Best Local Similarity   80.1%;  Pred. No. 0;
  Matches 2872;  Conservative    1;  Mismatches    3;  Indels  708;  Gaps    2;

Qy        455 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1311 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 1370

Qy        515 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1371 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 1430

Qy        575 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1431 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 1490

Qy        635 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 694
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1491 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 1550

Qy        695 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1551 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 1610

Qy        755 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 814
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1611 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 1670

Qy        815 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 874
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1671 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 1730

Qy        875 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 934
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1731 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 1790

Qy        935 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 994
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1791 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 1850

Qy        995 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1054
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1851 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1910

Qy       1055 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1911 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1970

Qy       1115 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 1174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1971 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 2030

Qy       1175 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 1234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2031 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 2090

Qy       1235 GATGA------------------------------------------------------- 1239
              |||||
Db       2091 GATGAAGGTAAGGCCTTGGCAGCCAGCAGAGGCTGGTGTGGGAGCCGCCCACCAGAGACG 2150

Qy       1240 ------------------------------------------------------------ 1239

Db       2151 ACACTCGGGGCTGTGTCTGGGCTGGTGCCTCTCCATCCTGGCCCCGACTTCTCTGTCAGG 2210

Qy       1240 ------------------------------------------------------------ 1239

Db       2211 AAAGTGGGGATGGACCCCATCTGCATACACGGCTTCTCATGGGTGTGGAACATCTCTGCT 2270

Qy       1240 ------------------------------------------------------------ 1239

Db       2271 TGCGGTTTCAGGAAGGCCTCTGGCTGCTCTAGGAGTCTGATCAGAGTCGTTGCCCCAGTT 2330

Qy       1240 ------------------------------------------------------------ 1239

Db       2331 TGACAGAAGGAAAGGCGGAGCTTATTCAAAGTCTAGAGGGAGTGGAGGAGTTAAGGCTGG 2390

Qy       1240 ------------------------------------------------------------ 1239

Db       2391 ATTTCAGATCTGCCTGGTTCCAGCCGCAGTGTGCCCTCTGCTCCCCCAACGACTTTCCAA 2450

Qy       1240 ------------------------------------------------------------ 1239

Db       2451 ATAATCTCACCAGCGCCTTCCAGCTCAGGCGTCCTAGAAGCGTCTTGAAGCCTATGGCCA 2510

Qy       1240 ------------------------------------------------------------ 1239

Db       2511 GCTGTCTTTGTGTTCCCTCTCACCCGCCTGTCCTCACAGCTGAGACTCCCAGGAAACCTT 2570

Qy       1240 ------------------------------------------------------------ 1239

Db       2571 CAGACTACCTTCCTCTGCCTTCAGCAAGGGGCGTTGCCCACATTCTCTGAGGGTCAGTGG 2630

Qy       1240 ------------------------------------------------------------ 1239

Db       2631 AAGAACCTAGACTCCCATTGCTAGAGGTAGAAAGGGGAAGGGTGCTGGGGAGCAGGGCTG 2690

Qy       1240 ------------------------------------------------------------ 1239

Db       2691 GTCCACAGCAGGTCTCGTGCAGCAGGTACCTGTGGTTCCGCCTTCTCATCTCCCTGAGAC 2750

Qy       1240 ---------------------------------------------------AGGCGTTCG 1248
                                                                 |||||||||
Db       2751 TGCTCCGACCCTTCCCTCCCAGGCTCTGTCTGATGGCCCCTCTCCCTCTGCAGGCGTTCG 2810

Qy       1249 GATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCAT 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2811 GATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCAT 2870

Qy       1309 GGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 1368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2871 GGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 2930

Qy       1369 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 1428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2931 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 2990

Qy       1429 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 1488
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2991 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 3050

Qy       1489 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 1548
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3051 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 3110

Qy       1549 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 1608
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3111 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 3170

Qy       1609 TAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 1668
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
```

```
Db       3171 TAATGGACACGTGGGTGYTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 3230

Qy       1669 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 1728
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3231 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 3290

Qy       1729 GGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTC 1788
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3291 GGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTC 3350

Qy       1789 CCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGC 1848
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3351 CCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGC 3410

Qy       1849 CTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT 1908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3411 CTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT 3470

Qy       1909 ATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG 1968
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3471 ATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG 3530

Qy       1969 GAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGG 2028
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db       3531 GAGGGCCTGCCTCACTGGGTCCCAGCTCCCTGCTCCTGTTAGCCCCATGGGGCTGCCGGG 3590

Qy       2029 CTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTG 2088
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3591 CTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTG 3650

Qy       2089 CTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTC 2148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3651 CTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTC 3710

Qy       2149 TAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGC 2208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3711 TAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGC 3770

Qy       2209 CAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG 2268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3771 CAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG 3830

Qy       2269 TTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAA 2328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3831 TTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAA 3890

Qy       2329 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCT 2388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3891 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCT 3950

Qy       2389 CTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTA 2446
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||  | |||
Db       3951 CTTGCATGGGAGTTTCTAGGATGAAACACTCCACCATGGGATTTGAACATATGAAAGTTA 4010

Qy       2447 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 2506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4011 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 4070

Qy       2507 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 2566
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4071 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 4130

Qy       2567 TTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGT 2626
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4131 TTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGT 4190

Qy       2627 AGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 2686
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4191 AGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 4250

Qy       2687 ATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCT 2746
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4251 ATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCT 4310

Qy       2747 TCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCA 2806
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4311 TCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCA 4370

Qy       2807 TCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGG 2866
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4371 TCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGG 4430

Qy       2867 GTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGT 2926
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4431 GTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGT 4490

Qy       2927 TCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 2986
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4491 TCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 4550

Qy       2987 AAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCC 3046
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4551 AAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCC 4610

Qy       3047 TGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATC 3106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4611 TGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATC 4670

Qy       3107 TCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCC 3166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4671 TCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCC 4730

Qy       3167 TGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATG 3226
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4731 TGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATG 4790

Qy       3227 TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAA 3286
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4791 TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAA 4850

Qy       3287 TGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAA 3330
              ||||||||||||||||||||||||||||||||||||||||||||
Db       4851 TGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAA 4894
```

```
RESULT 29
US-09-780-669-702
; Sequence 702, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C24
;  CURRENT APPLICATION NUMBER: US/09/780,669
;  CURRENT FILING DATE:  2001-02-09
;  NUMBER OF SEQ ID NOS:  943
```

```
; SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 702
;   LENGTH: 4894
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-780-669-702

  Query Match             62.8%;  Score 2142.8;  DB 10;  Length 4894;
  Best Local Similarity   80.1%;  Pred. No. 0;
  Matches 2872;  Conservative    1;  Mismatches    3;  Indels  708;  Gaps    2;

Qy        455 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1311 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 1370

Qy        515 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1371 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 1430

Qy        575 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1431 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 1490

Qy        635 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 694
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1491 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 1550

Qy        695 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1551 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 1610

Qy        755 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 814
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1611 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 1670

Qy        815 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 874
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1671 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 1730

Qy        875 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 934
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1731 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 1790

Qy        935 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 994
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1791 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 1850

Qy        995 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1054
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1851 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1910

Qy       1055 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1911 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1970

Qy       1115 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 1174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1971 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 2030

Qy       1175 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 1234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2031 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 2090

Qy       1235 GATGA------------------------------------------------------- 1239
              |||||
Db       2091 GATGAAGGTAAGGCCTTGGCAGCCAGCAGAGGCTGGTGTGGGAGCCGCCCACCAGAGACG 2150

Qy       1240 ------------------------------------------------------------ 1239

Db       2151 ACACTCGGGGCTGTGTCTGGGCTGGTGCCTCTCCATCCTGGCCCCGACTTCTCTGTCAGG 2210

Qy       1240 ------------------------------------------------------------ 1239

Db       2211 AAAGTGGGGATGGACCCCATCTGCATACACGGCTTCTCATGGGTGTGGAACATCTCTGCT 2270

Qy       1240 ------------------------------------------------------------ 1239

Db       2271 TGCGGTTTCAGGAAGGCCTCTGGCTGCTCTAGGAGTCTGATCAGAGTCGTTGCCCCAGTT 2330

Qy       1240 ------------------------------------------------------------ 1239

Db       2331 TGACAGAAGGAAAGGCGGAGCTTATTCAAAGTCTAGAGGGAGTGGAGGAGTTAAGGCTGG 2390

Qy       1240 ------------------------------------------------------------ 1239

Db       2391 ATTTCAGATCTGCCTGGTTCCAGCCGCAGTGTGCCCTCTGCTCCCCCAACGACTTTCCAA 2450

Qy       1240 ------------------------------------------------------------ 1239

Db       2451 ATAATCTCACCAGCGCCTTCCAGCTCAGGCGTCCTAGAAGCGTCTTGAAGCCTATGGCCA 2510

Qy       1240 ------------------------------------------------------------ 1239

Db       2511 GCTGTCTTTGTGTTCCCTCTCACCCGCCTGTCCTCACAGCTGAGACTCCCAGGAAACCTT 2570

Qy       1240 ------------------------------------------------------------ 1239

Db       2571 CAGACTACCTTCCTCTGCCTTCAGCAAGGGGCGTTGCCCACATTCTCTGAGGGTCAGTGG 2630

Qy       1240 ------------------------------------------------------------ 1239

Db       2631 AAGAACCTAGACTCCCATTGCTAGAGGTAGAAAGGGGAAGGGTGCTGGGGAGCAGGGCTG 2690

Qy       1240 ------------------------------------------------------------ 1239

Db       2691 GTCCACAGCAGGTCTCGTGCAGCAGGTACCTGTGGTTCCGCCTTCTCATCTCCCTGAGAC 2750

Qy       1240 ---------------------------------------------------AGGCGTTCG 1248
                                                                 |||||||||
Db       2751 TGCTCCGACCCTTCCCTCCCAGGCTCTGTCTGATGGCCCCTCTCCCTCTGCAGGCGTTCG 2810

Qy       1249 GATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCAT 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2811 GATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCAT 2870

Qy       1309 GGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 1368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2871 GGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 2930

Qy       1369 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 1428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2931 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 2990

Qy       1429 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 1488
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2991 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 3050

Qy       1489 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 1548
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3051 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 3110

Qy       1549 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 1608
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3111 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 3170

Qy       1609 TAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 1668
              |||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db       3171 TAATGGACACGTGGGTGYTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 3230

Qy       1669 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 1728
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3231 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 3290

Qy       1729 GGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTC 1788
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3291 GGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTC 3350
```

```
Qy      1789 CCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGC 1848
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3351 CCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGC 3410

Qy      1849 CTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT 1908
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3411 CTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT 3470

Qy      1909 ATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG 1968
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3471 ATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG 3530

Qy      1969 GAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGG 2028
             |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db      3531 GAGGGCCTGCCTCACTGGGTCCCAGCTCCCTGCTCCTGTTAGCCCCATGGGGCTGCCGGG 3590

Qy      2029 CTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTG 2088
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3591 CTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTG 3650

Qy      2089 CTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTC 2148
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3651 CTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTC 3710

Qy      2149 TAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGC 2208
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3711 TAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGC 3770

Qy      2209 CAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG 2268
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3771 CAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG 3830

Qy      2269 TTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAA 2328
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3831 TTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAA 3890

Qy      2329 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCT 2388
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3891 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCT 3950

Qy      2389 CTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTA 2446
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||  | |||
Db      3951 CTTGCATGGGAGTTTCTAGGATGAAACACTCCACCATGGGATTTGAACATATGAAAGTTA 4010

Qy      2447 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 2506
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4011 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 4070

Qy      2507 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 2566
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4071 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 4130

Qy      2567 TTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGT 2626
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4131 TTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGT 4190

Qy      2627 AGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 2686
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4191 AGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 4250

Qy      2687 ATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCT 2746
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4251 ATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCT 4310

Qy      2747 TCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCA 2806
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4311 TCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCA 4370

Qy      2807 TCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGG 2866
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4371 TCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGG 4430

Qy      2867 GTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGT 2926
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4431 GTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGT 4490

Qy      2927 TCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 2986
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4491 TCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 4550

Qy      2987 AAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCC 3046
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4551 AAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCC 4610

Qy      3047 TGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATC 3106
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4611 TGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATC 4670

Qy      3107 TCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCC 3166
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4671 TCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCC 4730

Qy      3167 TGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATG 3226
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4731 TGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATG 4790

Qy      3227 TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAA 3286
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4791 TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAA 4850

Qy      3287 TGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAA 3330
             ||||||||||||||||||||||||||||||||||||||||||||
Db      4851 TGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAA 4894
```

```
RESULT 30
US-09-822-827-702
; Sequence 702, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.534C1
;  CURRENT APPLICATION NUMBER: US/09/822,827
;  CURRENT FILING DATE:  2001-03-28
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 702
;   LENGTH: 4894
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-822-827-702

  Query Match             62.8%;  Score 2142.8;  DB 10;  Length 4894;
  Best Local Similarity   80.1%;  Pred. No. 0;
  Matches 2872;  Conservative    1;  Mismatches    3;  Indels  708;  Gaps    2;

Qy       455 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 514
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1311 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 1370

Qy       515 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 574
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1371 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 1430

Qy       575 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 634
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1431 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 1490

Qy       635 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 694
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1491 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 1550

Qy       695 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 754
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1551 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 1610

Qy       755 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 814
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1611 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 1670

Qy       815 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 874
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1671 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 1730

Qy       875 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 934
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1731 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 1790

Qy       935 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 994
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1791 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 1850

Qy       995 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1054
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1851 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1910

Qy      1055 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1114
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1911 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1970

Qy      1115 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 1174
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1971 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 2030

Qy      1175 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 1234
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2031 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 2090

Qy      1235 GATGA------------------------------------------------------- 1239
             |||||
Db      2091 GATGAAGGTAAGGCCTTGGCAGCCAGCAGAGGCTGGTGTGGGAGCCGCCCACCAGAGACG 2150

Qy      1240 ------------------------------------------------------------ 1239

Db      2151 ACACTCGGGGCTGTGTCTGGGCTGGTGCCTCTCCATCCTGGCCCCGACTTCTCTGTCAGG 2210

Qy      1240 ------------------------------------------------------------ 1239

Db      2211 AAAGTGGGGATGGACCCCATCTGCATACACGGCTTCTCATGGGTGTGGAACATCTCTGCT 2270

Qy      1240 ------------------------------------------------------------ 1239

Db      2271 TGCGGTTTCAGGAAGGCCTCTGGCTGCTCTAGGAGTCTGATCAGAGTCGTTGCCCCAGTT 2330

Qy      1240 ------------------------------------------------------------ 1239

Db      2331 TGACAGAAGGAAAGGCGGAGCTTATTCAAAGTCTAGAGGGAGTGGAGGAGTTAAGGCTGG 2390

Qy      1240 ------------------------------------------------------------ 1239

Db      2391 ATTTCAGATCTGCCTGGTTCCAGCCGCAGTGTGCCCTCTGCTCCCCCAACGACTTTCCAA 2450

Qy      1240 ------------------------------------------------------------ 1239

Db      2451 ATAATCTCACCAGCGCCTTCCAGCTCAGGCGTCCTAGAAGCGTCTTGAAGCCTATGGCCA 2510

Qy      1240 ------------------------------------------------------------ 1239

Db      2511 GCTGTCTTTGTGTTCCCTCTCACCCGCCTGTCCTCACAGCTGAGACTCCCAGGAAACCTT 2570

Qy      1240 ------------------------------------------------------------ 1239

Db      2571 CAGACTACCTTCCTCTGCCTTCAGCAAGGGGCGTTGCCCACATTCTCTGAGGGTCAGTGG 2630

Qy      1240 ------------------------------------------------------------ 1239

Db      2631 AAGAACCTAGACTCCCATTGCTAGAGGTAGAAAGGGGAAGGGTGCTGGGGAGCAGGGCTG 2690

Qy      1240 ------------------------------------------------------------ 1239

Db      2691 GTCCACAGCAGGTCTCGTGCAGCAGGTACCTGTGGTTCCGCCTTCTCATCTCCCTGAGAC 2750

Qy      1240 ---------------------------------------------------AGGCGTTCG 1248
                                                                |||||||||
Db      2751 TGCTCCGACCCTTCCCTCCCAGGCTCTGTCTGATGGCCCCTCTCCCTCTGCAGGCGTTCG 2810

Qy      1249 GATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCAT 1308
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2811 GATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCAT 2870

Qy      1309 GGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 1368
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2871 GGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 2930

Qy      1369 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 1428
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2931 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 2990

Qy      1429 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 1488
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2991 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 3050

Qy      1489 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 1548
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3051 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 3110

Qy      1549 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 1608
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3111 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 3170

Qy      1609 TAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 1668
             |||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db      3171 TAATGGACACGTGGGTGYTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 3230

Qy      1669 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 1728
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3231 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 3290

Qy      1729 GGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTC 1788
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3291 GGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTC 3350

Qy      1789 CCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGC 1848
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3351 CCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGC 3410

Qy      1849 CTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT 1908
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3411 CTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT 3470

Qy      1909 ATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG 1968
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3471 ATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG 3530

Qy      1969 GAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGG 2028
             |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db      3531 GAGGGCCTGCCTCACTGGGTCCCAGCTCCCTGCTCCTGTTAGCCCCATGGGGCTGCCGGG 3590

Qy      2029 CTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTG 2088
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3591 CTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTG 3650

Qy      2089 CTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTC 2148
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3651 CTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTC 3710

Qy      2149 TAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGC 2208
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3711 TAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGC 3770
```

```
Qy       2209 CAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG 2268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3771 CAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG 3830

Qy       2269 TTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTGGGAGCTGAATAAA 2328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3831 TTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTGGGAGCTGAATAAA 3890

Qy       2329 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCT 2388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3891 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCT 3950

Qy       2389 CTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTA 2446
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||   | |||
Db       3951 CTTGCATGGGAGTTTCTAGGATGAAACACTCCACCATGGGATTTGAACATATGAAAGTTA 4010

Qy       2447 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 2506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4011 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 4070

Qy       2507 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 2566
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4071 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 4130

Qy       2567 TTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGT 2626
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4131 TTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGT 4190

Qy       2627 AGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 2686
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4191 AGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 4250

Qy       2687 ATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCT 2746
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4251 ATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCT 4310

Qy       2747 TCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCA 2806
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4311 TCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCA 4370

Qy       2807 TCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGG 2866
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4371 TCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGG 4430

Qy       2867 GTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGT 2926
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4431 GTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGT 4490

Qy       2927 TCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 2986
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4491 TCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 4550

Qy       2987 AAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCC 3046
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4551 AAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCC 4610

Qy       3047 TGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATC 3106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4611 TGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATC 4670

Qy       3107 TCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCC 3166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4671 TCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCC 4730

Qy       3167 TGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATG 3226
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4731 TGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATG 4790

Qy       3227 TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAA 3286
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4791 TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAA 4850

Qy       3287 TGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAA 3330
              ||||||||||||||||||||||||||||||||||||||||||||
Db       4851 TGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAA 4894


RESULT 31
US-10-012-896-705
; Sequence 705, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Wantanabe, Yoshihiro
;  APPLICANT:  Meagher, Madeleine Joy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C27
;  CURRENT APPLICATION NUMBER: US/10/012,896
;  CURRENT FILING DATE:  2001-12-10
;  NUMBER OF SEQ ID NOS: 1011
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 705
;   LENGTH: 6976
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-012-896-705

  Query Match             53.2%;  Score 1815.8;  DB 9;  Length 6976;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 1841;  Conservative    0;  Mismatches   22;  Indels    2;  Gaps    1;

Qy       1470 CCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAG 1529
              |||   | ||  ||     ||| |  ||     |     ||||||||||||||||||||||
Db       5112 CCTGTCCTCTTCCCCTTTTCTTCACCCCTCTGCCTTAGGTGTTCCTGCCCAAATACCGAG 5171

Qy       1530 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTA 1589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5172 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTA 5231

Qy       1590 AGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCAC 1649
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5232 AGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCAC 5291

Qy       1650 CTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTG 1709
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5292 CTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTG 5351

Qy       1710 AGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGG 1769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5352 AGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGG 5411

Qy       1770 ATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGC 1829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       5412 ATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGC 5471

Qy       1830 TCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTT 1889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5472 TCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTT 5531

Qy       1890 ACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAA 1949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5532 ACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAA 5591

Qy       1950 CTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTA 2009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5592 CTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTA 5651

Qy       2010 GCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCT 2069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5652 GCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCT 5711

Qy       2070 CTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT 2129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5712 CTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT 5771

Qy       2130 CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTC 2189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5772 CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTC 5831

Qy       2190 TGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT 2249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5832 TGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT 5891

Qy       2250 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG 2309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5892 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG 5951

Qy       2310 TTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGC 2369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5952 TTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGC 6011

Qy       2370 AGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGA 2429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6012 AGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGA 6071

Qy       2430 TTTGAACATATGAC--TTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCA 2487
              |||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db       6072 TTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCA 6131

Qy       2488 GGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCA 2547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6132 GGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCA 6191

Qy       2548 GGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA 2607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6192 GGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA 6251

Qy       2608 CTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAA 2667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6252 CTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAA 6311

Qy       2668 TATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGG 2727
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6312 TATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGG 6371

Qy       2728 CTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGG 2787
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6372 CTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGG 6431

Qy       2788 ACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGG 2847
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6432 ACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGG 6491

Qy       2848 TGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCT 2907
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6492 TGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCT 6551

Qy       2908 TCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGG 2967
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6552 TCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGG 6611

Qy       2968 CTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTC 3027
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6612 CTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTC 6671

Qy       3028 CCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 3087
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6672 CCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 6731

Qy       3088 TCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACA 3147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6732 TCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACA 6791

Qy       3148 GAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTA 3207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6792 GAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTA 6851

Qy       3208 AGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 3267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6852 AGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 6911

Qy       3268 GTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAA 3327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6912 GTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAA 6971

Qy       3328 AAAAA 3332
              |||||
Db       6972 AAAAA 6976

RESULT 32
US-09-895-793-705
; Sequence 705, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.534C2
;  CURRENT APPLICATION NUMBER: US/09/895,793
;  CURRENT FILING DATE:  2001-06-29
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 705
;   LENGTH: 6976
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-895-793-705
```

```
Query Match             53.2%;  Score 1815.8;  DB 9;  Length 6976;
Best Local Similarity   98.7%;  Pred. No. 0;
Matches 1841;  Conservative    0;  Mismatches   22;  Indels    2;  Gaps    1;

Qy       1470 CCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAG 1529
              |||   | ||  ||    ||| |  ||    |    ||||||||||||||||||||||||
Db       5112 CCTGTCCTCTTCCCCTTTTCTTCACCCCTCTGCCTTAGGTGTTCCTGCCCAAATACCGAG 5171

Qy       1530 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTA 1589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5172 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTA 5231

Qy       1590 AGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCAC 1649
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5232 AGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCAC 5291

Qy       1650 CTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTG 1709
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5292 CTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTG 5351

Qy       1710 AGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGG 1769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5352 AGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGG 5411

Qy       1770 ATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGC 1829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5412 ATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGC 5471

Qy       1830 TCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTT 1889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5472 TCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTT 5531

Qy       1890 ACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAA 1949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5532 ACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAA 5591

Qy       1950 CTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTA 2009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5592 CTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTA 5651

Qy       2010 GCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCT 2069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5652 GCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCT 5711

Qy       2070 CTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT 2129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5712 CTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT 5771

Qy       2130 CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTC 2189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5772 CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTC 5831

Qy       2190 TGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT 2249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5832 TGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT 5891

Qy       2250 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG 2309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5892 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG 5951

Qy       2310 TTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGC 2369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5952 TTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGC 6011

Qy       2370 AGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGA 2429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6012 AGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGA 6071

Qy       2430 TTTGAACATATGAC--TTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCA 2487
              |||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db       6072 TTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCA 6131

Qy       2488 GGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCA 2547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6132 GGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCA 6191

Qy       2548 GGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA 2607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6192 GGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA 6251

Qy       2608 CTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAA 2667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6252 CTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAA 6311

Qy       2668 TATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGG 2727
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6312 TATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGG 6371

Qy       2728 CTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGG 2787
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6372 CTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGG 6431

Qy       2788 ACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGG 2847
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6432 ACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGG 6491

Qy       2848 TGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCT 2907
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6492 TGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCT 6551

Qy       2908 TCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGG 2967
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6552 TCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGG 6611

Qy       2968 CTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTC 3027
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6612 CTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTC 6671

Qy       3028 CCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 3087
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6672 CCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 6731

Qy       3088 TCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACA 3147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6732 TCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACA 6791

Qy       3148 GAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTA 3207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6792 GAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTA 6851

Qy       3208 AGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 3267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6852 AGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 6911

Qy       3268 GTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAA 3327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6912 GTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAA 6971

Qy       3328 AAAAA 3332
              |||||
Db       6972 AAAAA 6976


RESULT 33
US-09-895-814-705
; Sequence 705, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
```

```
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C26
;  CURRENT APPLICATION NUMBER: US/09/895,814
;  CURRENT FILING DATE:  2001-06-29
;  NUMBER OF SEQ ID NOS: 990
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 705
;   LENGTH: 6976
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-895-814-705

  Query Match             53.2%;  Score 1815.8;  DB 9;  Length 6976;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 1841;  Conservative    0;  Mismatches   22;  Indels    2;  Gaps    1;

Qy       1470 CCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAG 1529
              |||   | ||  ||   |||  |  ||     |   |||||||||||||||||||||||
Db       5112 CCTGTCCTCTTCCCCTTTTCTTCACCCCTCTGCCTTAGGTGTTCCTGCCCAAATACCGAG 5171

Qy       1530 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTA 1589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5172 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTA 5231

Qy       1590 AGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCAC 1649
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5232 AGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCAC 5291

Qy       1650 CTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTG 1709
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5292 CTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTG 5351

Qy       1710 AGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGG 1769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5352 AGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGG 5411

Qy       1770 ATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGC 1829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5412 ATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGC 5471

Qy       1830 TCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTT 1889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5472 TCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTT 5531

Qy       1890 ACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAA 1949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5532 ACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAA 5591

Qy       1950 CTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTA 2009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5592 CTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTA 5651

Qy       2010 GCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCT 2069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5652 GCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCT 5711

Qy       2070 CTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT 2129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5712 CTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT 5771

Qy       2130 CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTC 2189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5772 CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTC 5831

Qy       2190 TGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT 2249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5832 TGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT 5891

Qy       2250 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG 2309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5892 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG 5951

Qy       2310 TTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGC 2369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5952 TTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGC 6011

Qy       2370 AGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGA 2429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6012 AGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGA 6071

Qy       2430 TTTGAACATATGAC--TTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCA 2487
              ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db       6072 TTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCA 6131

Qy       2488 GGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCA 2547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6132 GGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCA 6191

Qy       2548 GGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA 2607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6192 GGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA 6251

Qy       2608 CTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAA 2667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6252 CTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAA 6311

Qy       2668 TATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGG 2727
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6312 TATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGG 6371

Qy       2728 CTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGG 2787
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6372 CTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGG 6431

Qy       2788 ACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGG 2847
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6432 ACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGG 6491

Qy       2848 TGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCT 2907
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6492 TGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCT 6551

Qy       2908 TCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGG 2967
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6552 TCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGG 6611

Qy       2968 CTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTC 3027
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6612 CTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTC 6671

Qy       3028 CCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 3087
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6672 CCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 6731

Qy       3088 TCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACA 3147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6732 TCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACA 6791

Qy       3148 GAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTA 3207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       6792 GAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTA 6851

Qy       3208 AGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 3267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6852 AGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 6911

Qy       3268 GTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAA 3327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6912 GTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAA 6971

Qy       3328 AAAAA 3332
              |||||
Db       6972 AAAAA 6976
```

```
RESULT 34
US-09-759-143-705
; Sequence 705, Application US/09759143
; Patent No. US20020022248A1
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C23
;  CURRENT APPLICATION NUMBER: US/09/759,143
;  CURRENT FILING DATE:  2001-01-12
;  NUMBER OF SEQ ID NOS:  934
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 705
;   LENGTH: 6976
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-759-143-705

  Query Match             53.2%;  Score 1815.8;  DB 10;  Length 6976;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 1841;  Conservative    0;  Mismatches   22;  Indels    2;  Gaps    1;

Qy       1470 CCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAG 1529
              |||   | ||  || |  ||| |  ||    |    ||||||||||||||||||||||||
Db       5112 CCTGTCCTCTTCCCCTTTTCTTCACCCCTCTGCCTTAGGTGTTCCTGCCCAAATACCGAG 5171

Qy       1530 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTA 1589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5172 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTA 5231

Qy       1590 AGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCAC 1649
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5232 AGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCAC 5291

Qy       1650 CTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTG 1709
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5292 CTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTG 5351

Qy       1710 AGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGG 1769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5352 AGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGG 5411

Qy       1770 ATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGC 1829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5412 ATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGC 5471

Qy       1830 TCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTT 1889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5472 TCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTT 5531

Qy       1890 ACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAA 1949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5532 ACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAA 5591

Qy       1950 CTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTA 2009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5592 CTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTA 5651

Qy       2010 GCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCT 2069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5652 GCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCT 5711

Qy       2070 CTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT 2129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5712 CTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT 5771

Qy       2130 CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTC 2189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5772 CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTC 5831

Qy       2190 TGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT 2249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5832 TGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT 5891

Qy       2250 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG 2309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5892 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG 5951

Qy       2310 TTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGC 2369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5952 TTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGC 6011

Qy       2370 AGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGA 2429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6012 AGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGA 6071

Qy       2430 TTTGAACATATGAC--TTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCA 2487
              |||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db       6072 TTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCA 6131

Qy       2488 GGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCA 2547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6132 GGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCA 6191

Qy       2548 GGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA 2607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6192 GGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA 6251

Qy       2608 CTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAA 2667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6252 CTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAA 6311

Qy       2668 TATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGG 2727
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6312 TATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGG 6371

Qy       2728 CTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGG 2787
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6372 CTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGG 6431

Qy       2788 ACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGG 2847
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6432 ACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGG 6491
```

```
Qy       2848 TGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCT 2907
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6492 TGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCT 6551

Qy       2908 TCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGG 2967
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6552 TCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGG 6611

Qy       2968 CTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTC 3027
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6612 CTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTC 6671

Qy       3028 CCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 3087
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6672 CCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 6731

Qy       3088 TCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACA 3147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6732 TCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACA 6791

Qy       3148 GAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTA 3207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6792 GAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTA 6851

Qy       3208 AGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 3267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6852 AGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 6911

Qy       3268 GTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAA 3327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6912 GTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAA 6971

Qy       3328 AAAAA 3332
              |||||
Db       6972 AAAAA 6976
```

```
RESULT 35
US-09-780-669-705
; Sequence 705, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C24
;  CURRENT APPLICATION NUMBER: US/09/780,669
;  CURRENT FILING DATE:  2001-02-09
;  NUMBER OF SEQ ID NOS:  943
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 705
;   LENGTH: 6976
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

```
US-09-780-669-705

  Query Match             53.2%;  Score 1815.8;  DB 10;  Length 6976;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 1841;  Conservative    0;  Mismatches   22;  Indels    2;  Gaps    1;

Qy       1470 CCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAG 1529
              |||   | ||  ||    ||| |  ||    |    ||||||||||||||||||||||||
Db       5112 CCTGTCCTCTTCCCCTTTTCTTCACCCCTCTGCCTTAGGTGTTCCTGCCCAAATACCGAG 5171

Qy       1530 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTA 1589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5172 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTA 5231

Qy       1590 AGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCAC 1649
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5232 AGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCAC 5291

Qy       1650 CTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTG 1709
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5292 CTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTG 5351

Qy       1710 AGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGG 1769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5352 AGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGG 5411

Qy       1770 ATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGC 1829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5412 ATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGC 5471

Qy       1830 TCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTT 1889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5472 TCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTT 5531

Qy       1890 ACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAA 1949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5532 ACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAA 5591

Qy       1950 CTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTA 2009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5592 CTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTA 5651

Qy       2010 GCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCT 2069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5652 GCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCT 5711

Qy       2070 CTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT 2129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5712 CTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT 5771

Qy       2130 CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTC 2189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5772 CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTC 5831

Qy       2190 TGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT 2249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5832 TGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT 5891

Qy       2250 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG 2309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5892 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG 5951

Qy       2310 TTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGC 2369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5952 TTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGC 6011

Qy       2370 AGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGA 2429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6012 AGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGA 6071

Qy       2430 TTTGAACATATGAC--TTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCA 2487
              |||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db       6072 TTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCA 6131
```

```
Qy       2488 GGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCA 2547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6132 GGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCA 6191

Qy       2548 GGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA 2607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6192 GGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA 6251

Qy       2608 CTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAA 2667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6252 CTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAA 6311

Qy       2668 TATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGG 2727
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6312 TATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGG 6371

Qy       2728 CTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGG 2787
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6372 CTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGG 6431

Qy       2788 ACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGG 2847
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6432 ACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGG 6491

Qy       2848 TGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCT 2907
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6492 TGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCT 6551

Qy       2908 TCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGG 2967
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6552 TCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGG 6611

Qy       2968 CTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTC 3027
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6612 CTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTC 6671

Qy       3028 CCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 3087
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6672 CCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 6731

Qy       3088 TCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACA 3147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6732 TCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACA 6791

Qy       3148 GAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGGTTTA 3207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6792 GAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGGTTTA 6851

Qy       3208 AGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 3267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6852 AGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 6911

Qy       3268 GTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAA 3327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6912 GTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAA 6971

Qy       3328 AAAAA 3332
              |||||
Db       6972 AAAAA 6976
```

RESULT 36
US-09-822-827-705
; Sequence 705, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 705
;   LENGTH: 6976
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-822-827-705

```
  Query Match             53.2%;  Score 1815.8;  DB 10;  Length 6976;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 1841;  Conservative    0;  Mismatches   22;  Indels    2;  Gaps    1;

Qy       1470 CCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAG 1529
               |||   | ||  ||    ||| |  ||     |     ||||||||||||||||||||||
Db       5112 CCTGTCCTCTTCCCCTTTTCTTCACCCCTCTGCCTTAGGTGTTCCTGCCCAAATACCGAG 5171

Qy       1530 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTA 1589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5172 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTA 5231

Qy       1590 AGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCAC 1649
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5232 AGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCAC 5291

Qy       1650 CTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTG 1709
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5292 CTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTG 5351

Qy       1710 AGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGG 1769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5352 AGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGG 5411

Qy       1770 ATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGC 1829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5412 ATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGC 5471

Qy       1830 TCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTT 1889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5472 TCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTT 5531

Qy       1890 ACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAA 1949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5532 ACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAA 5591

Qy       1950 CTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTA 2009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5592 CTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTA 5651

Qy       2010 GCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCT 2069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5652 GCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCT 5711

Qy       2070 CTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT 2129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5712 CTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT 5771

Qy       2130 CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTC 2189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5772 CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTC 5831

Qy       2190 TGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT 2249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5832 TGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT 5891

Qy       2250 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG 2309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5892 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG 5951

Qy       2310 TTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGC 2369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5952 TTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGC 6011

Qy       2370 AGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGA 2429
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6012 AGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGA 6071

Qy        2430 TTTGAACATATGAC--TTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCA 2487
               |||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        6072 TTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCA 6131

Qy        2488 GGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCA 2547
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6132 GGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCA 6191

Qy        2548 GGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA 2607
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6192 GGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA 6251

Qy        2608 CTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAA 2667
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6252 CTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAA 6311

Qy        2668 TATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGG 2727
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6312 TATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGG 6371

Qy        2728 CTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGG 2787
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6372 CTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGG 6431

Qy        2788 ACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGG 2847
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6432 ACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGG 6491

Qy        2848 TGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCT 2907
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6492 TGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCT 6551

Qy        2908 TCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGG 2967
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6552 TCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGG 6611

Qy        2968 CTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTC 3027
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6612 CTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTC 6671

Qy        3028 CCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 3087
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6672 CCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 6731

Qy        3088 TCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACA 3147
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6732 TCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACA 6791

Qy        3148 GAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTA 3207
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6792 GAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTA 6851

Qy        3208 AGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 3267
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6852 AGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 6911

Qy        3268 GTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAA 3327
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6912 GTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAA 6971

Qy        3328 AAAAA 3332
               |||||
Db        6972 AAAAA 6976
```

```
RESULT 37
US-10-012-896-851
; Sequence 851, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Wantanabe, Yoshihiro
;  APPLICANT:  Meagher, Madeleine Joy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C27
;  CURRENT APPLICATION NUMBER: US/10/012,896
;  CURRENT FILING DATE:  2001-12-10
;  NUMBER OF SEQ ID NOS: 1011
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 851
;   LENGTH: 1203
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-012-896-851

  Query Match             23.3%;  Score 794.6;  DB 9;  Length 1203;
  Best Local Similarity   99.5%;  Pred. No. 4.5e-142;
  Matches  797;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         379 GGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTT 438
               ||||| |  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         402 GGCCGAATTCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTT 461

Qy         439 CATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGG 498
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         462 CATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGG 521

Qy         499 CTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACT 558
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         522 CTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACT 581

Qy         559 GTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCT 618
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         582 GTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCT 641

Qy         619 GCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCT 678
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         642 GCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCT 701

Qy         679 GGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCG 738
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         702 GGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCG 761

Qy         739 GGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGG 798
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         762 GGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGG 821

Qy         799 CTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCT 858
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         822 CTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCT 881
```

```
Qy        859 GGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGC 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        882 GGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGC 941

Qy        919 AGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCT 978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        942 AGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCT 1001

Qy        979 GTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAA 1038
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1002 GTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAA 1061

Qy       1039 CCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCG 1098
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1062 CCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCG 1121

Qy       1099 CCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTA 1158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1122 CCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTA 1181

Qy       1159 CACGGATTTCGTGGGCGAGGG 1179
              ||||||||||||||||||| |
Db       1182 CACGGATTTCGTGGGCGAGTG 1202
```

```
RESULT 38
US-09-895-793-851
; Sequence 851, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT;  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.534C2
;  CURRENT APPLICATION NUMBER: US/09/895,793
;  CURRENT FILING DATE:  2001-06-29
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 851
;   LENGTH: 1203
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-895-793-851

  Query Match             23.3%;  Score 794.6;  DB 9;  Length 1203;
  Best Local Similarity   99.5%;  Pred. No. 4.5e-142;
  Matches  797;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        379 GGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTT 438
              ||||| |  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 GGCCGAATTCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTT 461

Qy        439 CATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGG 498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 CATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGG 521

Qy        499 CTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACT 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 CTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACT 581

Qy        559 GTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCT 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 GTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCT 641

Qy        619 GCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCT 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        642 GCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCT 701

Qy        679 GGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCG 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        702 GGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCG 761

Qy        739 GGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGG 798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        762 GGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGG 821

Qy        799 CTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCT 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        822 CTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCT 881

Qy        859 GGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGC 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        882 GGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGC 941

Qy        919 AGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCT 978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        942 AGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCT 1001

Qy        979 GTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAA 1038
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1002 GTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAA 1061

Qy       1039 CCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCG 1098
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1062 CCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCG 1121

Qy       1099 CCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTA 1158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1122 CCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTA 1181

Qy       1159 CACGGATTTCGTGGGCGAGGG 1179
              ||||||||||||||||||| |
Db       1182 CACGGATTTCGTGGGCGAGTG 1202
```

```
RESULT 39
US-09-895-814-851
; Sequence 851, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
```

```
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C26
;  CURRENT APPLICATION NUMBER: US/09/895,814
;  CURRENT FILING DATE:  2001-06-29
;  NUMBER OF SEQ ID NOS: 990
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 851
;   LENGTH: 1203
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-895-814-851

  Query Match             23.3%;  Score 794.6;  DB 9;  Length 1203;
  Best Local Similarity   99.5%;  Pred. No. 4.5e-142;
  Matches  797;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        379 GGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTT 438
              ||||| |  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 GGCCGAATTCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTT 461

Qy        439 CATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGG 498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 CATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGG 521

Qy        499 CTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACT 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 CTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACT 581

Qy        559 GTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCT 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 GTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCT 641

Qy        619 GCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCT 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        642 GCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCT 701

Qy        679 GGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCG 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        702 GGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCG 761

Qy        739 GGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGG 798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        762 GGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGG 821

Qy        799 CTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCT 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        822 CTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCT 881

Qy        859 GGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGC 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        882 GGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGC 941

Qy        919 AGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCT 978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        942 AGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCT 1001

Qy        979 GTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAA 1038
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1002 GTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAA 1061

Qy       1039 CCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCG 1098
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1062 CCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCG 1121

Qy       1099 CCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTA 1158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1122 CCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTA 1181

Qy       1159 CACGGATTTCGTGGCGAGGG 1179
              |||||||||||||||||| |
Db       1182 CACGGATTTCGTGGCGAGTG 1202


RESULT 40
US-09-759-143-851
; Sequence 851, Application US/09759143
; Patent No. US20020022248A1
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C23
;  CURRENT APPLICATION NUMBER: US/09/759,143
;  CURRENT FILING DATE:  2001-01-12
;  NUMBER OF SEQ ID NOS:  934
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 851
;   LENGTH: 1203
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-759-143-851

  Query Match             23.3%;  Score 794.6;  DB 10;  Length 1203;
  Best Local Similarity   99.5%;  Pred. No. 4.5e-142;
  Matches  797;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        379 GGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTT 438
              ||||| |  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 GGCCGAATTCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTT 461

Qy        439 CATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGG 498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 CATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGG 521

Qy        499 CTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACT 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 CTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACT 581

Qy        559 GTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCT 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 GTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCT 641

Qy        619 GCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCT 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        642 GCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCT 701

Qy        679 GGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCG 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        702 GGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCG 761

Qy        739 GGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGG 798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        762 GGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGG 821

Qy        799 CTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCT 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        822 CTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCT 881

Qy        859 GGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGC 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        882 GGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGC 941

Qy        919 AGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCT 978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        942 AGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCT 1001

Qy        979 GTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAA 1038
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1002 GTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAA 1061

Qy       1039 CCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCG 1098
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1062 CCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCG 1121

Qy       1099 CCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTA 1158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1122 CCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTA 1181

Qy       1159 CACGGATTTCGTGGGCGAGGG 1179
              ||||||||||||||||||| |
Db       1182 CACGGATTTCGTGGGCGAGTG 1202
```

RESULT 41
US-09-780-669-851
; Sequence 851, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780,669
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 851
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-780-669-851

```
  Query Match            23.3%;  Score 794.6;  DB 10;  Length 1203;
  Best Local Similarity  99.5%;  Pred. No. 4.5e-142;
  Matches 797;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        379 GGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTT 438
              ||||| |  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 GGCCGAATTCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTT 461

Qy        439 CATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGG 498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 CATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGG 521

Qy        499 CTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACT 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 CTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACT 581

Qy        559 GTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCT 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 GTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCT 641

Qy        619 GCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCT 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        642 GCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCT 701

Qy        679 GGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCG 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        702 GGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCG 761

Qy        739 GGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGG 798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        762 GGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGG 821

Qy        799 CTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCT 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        822 CTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCT 881

Qy        859 GGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGC 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        882 GGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGC 941

Qy        919 AGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCT 978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        942 AGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCT 1001

Qy        979 GTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAA 1038
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1002 GTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAA 1061

Qy       1039 CCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCG 1098
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1062 CCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCG 1121

Qy       1099 CCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTA 1158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1122 CCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTA 1181

Qy       1159 CACGGATTTCGTGGGCGAGGG 1179
              ||||||||||||||||||| |
Db       1182 CACGGATTTCGTGGGCGAGTG 1202
```

RESULT 42
US-09-822-827-851
; Sequence 851, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 851
; LENGTH: 1203
; TYPE: DNA

```
;   ORGANISM: Homo sapiens
US-09-822-827-851

  Query Match             23.3%;  Score 794.6;  DB 10;  Length 1203;
  Best Local Similarity   99.5%;  Pred. No. 4.5e-142;
  Matches  797;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        379 GGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTT 438
              ||||| |  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 GGCCGAATTCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTT 461

Qy        439 CATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGG 498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 CATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGG 521

Qy        499 CTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACT 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 CTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACT 581

Qy        559 GTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCT 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 GTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCT 641

Qy        619 GCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCT 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        642 GCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCT 701

Qy        679 GGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCG 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        702 GGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCG 761

Qy        739 GGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGG 798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        762 GGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGG 821

Qy        799 CTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCT 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        822 CTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCT 881

Qy        859 GGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGC 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        882 GGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGC 941

Qy        919 AGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCT 978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        942 AGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCT 1001

Qy        979 GTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAA 1038
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1002 GTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAA 1061

Qy       1039 CCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCG 1098
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1062 CCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCG 1121

Qy       1099 CCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTA 1158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1122 CCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTA 1181

Qy       1159 CACGGATTTCGTGGGCGAGGG 1179
              ||||||||||||||||||| |
Db       1182 CACGGATTTCGTGGGCGAGTG 1202

RESULT 43
US-09-232-880-10
; Sequence 10, Application US/09232880
; Publication No. US20020182596A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT: Dillon, Davin C.
;  APPLICANT: Mitcham, Jennifer Lynn
;  TITLE OF INVENTION: COMPOUNDS FOR IMMUNODIAGNOSIS OF
;  TITLE OF INVENTION: PROSTATE CANCER AND METHODS FOR THEIR USE
;  FILE REFERENCE: 210121.428C6
;  CURRENT APPLICATION NUMBER: US/09/232,880
;  CURRENT FILING DATE: 1999-01-15
;  NUMBER OF SEQ ID NOS: 338
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10
;   LENGTH: 789
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(789)
;   OTHER INFORMATION: n = A,T,C or G
US-09-232-880-10

  Query Match             19.7%;  Score 673.4;  DB 9;  Length 789;
  Best Local Similarity   94.2%;  Pred. No. 4.8e-119;
  Matches  745;  Conservative    0;  Mismatches   40;  Indels    6;  Gaps    5;

Qy       1341 CAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCC 1400
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGTCTATNTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCC 60

Qy       1401 ACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGC 1460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGC 120

Qy       1461 AGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCA 1520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCA 180

Qy       1521 AATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC 1580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC 240

Qy       1581 CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCC 1640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCC 300

Qy       1641 TGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGG 1700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGG 360

Qy       1701 TGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCG 1760
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        361 TGGTGGGTGAGCCCACCGANGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCG 420

Qy       1761 CCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGCTCCA 1820
              ||||||||||||||| ||||||||||||||| ||||||||||||||||||||||||||||
Db        421 CCATCCTGGATAGTG-CTTCCTGCTGTCCCANGTGGCCCCATCCCTGTTTATGGCTCCA 479

Qy       1821 TTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGG 1880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 TTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGG 539

Qy       1881 TCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAG 1940
              || ||||||||||||||||||||||| |||||||||||| ||| ||||||||||||||||
Db        540 TC-CCATTTACTTTGCTACACAGGTANTATTTGACAAGAACGANTTGGCCAAATACTCAG 598

Qy       1941 CGTAGAAAACTTCCAGCA--CATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCC 1998
              |||  ||||  |||||||       |  ||||||||  ||||||||||||||||  |||  |||||
Db        599 CGTTAAAAAATTCCAGCAACATTGGGGGTGGAAGGCCTGCCTCACTGGGT-CCAACTCCC 657

Qy       1999 CGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGT 2058
              ||||||||||| ||||||||||||||||  ||||||||||| ||||||||||||||||| |
Db        658 CGCTCCTGTTAACCCCATGGGGCTGCCGGCTTGGCCGCCAATTTCTGTTGCTGCCAAANT 717

Qy       2059 AATGTGGCTCTCTGCTGCCACCCTGTGCT-GCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2117
              ||||||||||||||||||||     |||| |||||  |||| ||            |||||
Db        718 NATGTGGCTCTCTGCTGCCACCTGTTGCTGGCTGAAGTGCNTACNGCNCANCTNGGGGGG 777

Qy       2118 TGGGGCGTCCC 2128
```

```
                | ||| || ||
Db        778 TNGGGNGTTCC 788


RESULT 44
US-10-012-896-10
; Sequence 10, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Wantanabe, Yoshihiro
;  APPLICANT:  Meagher, Madeleine Joy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C27
;  CURRENT APPLICATION NUMBER: US/10/012,896
;  CURRENT FILING DATE:  2001-12-10
;  NUMBER OF SEQ ID NOS: 1011
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10
;   LENGTH: 789
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 9, 380, 451, 565, 582, 716, 718, 758, 762, 765, 768, 771,
;   LOCATION: 779, 783
;   OTHER INFORMATION: n = A,T,C or G
US-10-012-896-10

  Query Match             19.7%;  Score 673.4;  DB 9;  Length 789;
  Best Local Similarity   94.2%;  Pred. No. 4.8e-119;
  Matches  745;  Conservative    0;  Mismatches   40;  Indels    6;  Gaps    5;

Qy       1341 CAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCC 1400
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGTCTATNTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCC 60

Qy       1401 ACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGC 1460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGC 120

Qy       1461 AGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCA 1520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCA 180

Qy       1521 AATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC 1580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC 240

Qy       1581 CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCC 1640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCC 300

Qy       1641 TGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGG 1700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGG 360

Qy       1701 TGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCG 1760
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGTGGGTGAGCCCACCGANGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCG 420

Qy       1761 CCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCA 1820
              |||||||||||||| |||||||||||||||| ||||||||||||||||||||||||||||
Db        421 CCATCCTGGATAGTG-CTTCCTGCTGTCCCANGTGGCCCCATCCCTGTTTATGGGCTCCA 479

Qy       1821 TTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGG 1880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 TTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGG 539

Qy       1881 TCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAG 1940
              || ||||||||||||||||||||| |||||||||||| ||| |||||||||||||||||
Db        540 TC-CCATTTACTTTGCTACACAGGTANTATTTGACAAGAACGANTTGGCCAAATACTCAG 598

Qy       1941 CGTAGAAAACTTCCAGCA--CATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCC 1998
              |||   |||| ||||||||     | |||||||| ||||||||||||||||| ||| |||||
Db        599 CGTTAAAAAATTCCAGCAACATTGGGGGTGGAAGGCCTGCCTCACTGGGT-CCAACTCCC 657

Qy       1999 CGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGT 2058
              ||||||||||| |||||||||||||||||| ||||||||| |||||||||||||||| |
Db        658 CGCTCCTGTTAACCCCATGGGGCTGCCGGCTTGGCCGCCAATTTCTGTTGCTGCCAAANT 717

Qy       2059 AATGTGGCTCTCTGCTGCCACCCTGTGCT-GCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2117
               ||||||||||||||||||||||    |||| ||||| |||| ||          |||||
Db        718 NATGTGGCTCTCTGCTGCCACCTGTTGCTGGCTGAAGTGCNTACNGCNCANCTNGGGGGG 777

Qy       2118 TGGGGCGTCCC 2128
               | ||| || ||
Db        778 TNGGGNGTTCC 788


RESULT 45
US-09-895-793-10
; Sequence 10, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.534C2
;  CURRENT APPLICATION NUMBER: US/09/895,793
;  CURRENT FILING DATE:  2001-06-29
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
```

```
; SEQ ID NO 10
;   LENGTH: 789
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(789)
;   OTHER INFORMATION: n = A,T,C or G
US-09-895-793-10

  Query Match             19.7%;  Score 673.4;  DB 9;  Length 789;
  Best Local Similarity   94.2%;  Pred. No. 4.8e-119;
  Matches  745;  Conservative    0;  Mismatches   40;  Indels    6;  Gaps    5;

Qy       1341 CAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCC 1400
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGTCTATNTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCC 60

Qy       1401 ACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGC 1460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGC 120

Qy       1461 AGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCA 1520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCA 180

Qy       1521 AATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC 1580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC 240

Qy       1581 CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCC 1640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCC 300

Qy       1641 TGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGG 1700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGG 360

Qy       1701 TGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCG 1760
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        361 TGGTGGGTGAGCCCACCGANGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCG 420

Qy       1761 CCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCA 1820
              ||||||||||||||| |||||||||||||||| ||||||||||||||||||||||||||||
Db        421 CCATCCTGGATAGTG-CTTCCTGCTGTCCCANGTGGCCCCATCCCTGTTTATGGGCTCCA 479

Qy       1821 TTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGG 1880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 TTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGG 539

Qy       1881 TCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAG 1940
              || ||||||||||||||||||||||| |||||||||||| ||| ||||||||||||||||
Db        540 TC-CCATTTACTTTGCTACACAGGTANTATTTGACAAGAACGANTTGGCCAAATACTCAG 598

Qy       1941 CGTAGAAAACTTCCAGCA--CATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCC 1998
              |||  |||| ||||||||    | |||||||| ||||||||||||||||| ||| |||||
Db        599 CGTTAAAAAATTCCAGCAACATTGGGGGTGGAAGGCCTGCCTCACTGGGT-CCAACTCCC 657

Qy       1999 CGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGT 2058
              ||||||||||| |||||||||||||||||  ||||||||| ||||||||||||||||| |
Db        658 CGCTCCTGTTAACCCCATGGGGCTGCCGGCTTGGCCGCCAATTTCTGTTGCTGCCAAANT 717

Qy       2059 AATGTGGCTCTCTGCTGCCACCCTGTGCT-GCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2117
               |||||||||||||||||||||   |||| ||||| |||| ||           ||||| 
Db        718 NATGTGGCTCTCTGCTGCCACCTGTTGCTGGCTGAAGTGCNTACNGCNCANCTNGGGGGG 777

Qy       2118 TGGGGCGTCCC 2128
              | ||| || ||
Db        778 TNGGGNGTTCC 788

Search completed: June  9, 2003, 22:23:03
Job time : 609 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:         June  9, 2003, 19:43:44 ; Search time 85.5 Seconds
                                          (without alignments)
                                          10628.880 Million cell updates/sec

Title:          US-09-759-143-110
Perfect score:  6418
Sequence:       1 gggaaccagcctgcacgcgc..........aaaaaaataaaaaaaaaaaa 3410

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters:      1816940

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09759143/runat_05062003_111318_23032/app_query.fasta_1.3591
-DB=A_Geneseq_101002 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=45
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09759143_@CGN_1_1_119_@runat_05062003_111318_23032 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      A_Geneseq_101002:*
                1:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
                2:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
                3:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
                4:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
                5:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
                6:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
                7:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
                8:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
                9:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
                10:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
                11:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
                12:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
                13:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
                14:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
                15:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
                16:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
                17:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
                18:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
                19:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
                20:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
                21:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
                22:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
                23:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2861	44.6	553	19	AAW71869	Amino acid encoded
	2	2861	44.6	553	19	AAW69385	Prostate tumour sp
	3	2861	44.6	553	21	AAB28527	Protein encoded by
	4	2861	44.6	553	21	AAY82002	Human immunogenic
	5	2861	44.6	553	22	AAU69763	Human prostate cDN
	6	2861	44.6	553	22	AAU04961	Human prostate tum
	7	2861	44.6	553	22	AAM01117	Human prostate-spe
	8	2861	44.6	553	22	AAG99002	Human prostate-spe
	9	2861	44.6	553	22	AAG62150	Human P501S invent
	10	2861	44.6	553	22	AAB74800	Prostate tumour an
	11	2861	44.6	553	23	ABB77575	Human mast cell re
	12	2861	44.6	553	23	ABG61900	Prostate cancer-as
	13	2861	44.6	553	23	ABB95222	Human L1-12 protei
	14	2861	44.6	553	23	AAU10324	Human PROST 03. H
	15	2861	44.6	553	23	AAU82643	Human breast tumou
	16	2861	44.6	1079	22	AAB74830	Prostate tumour an
	17	2601	40.5	710	23	AAM50661	Thioredoxin-ubiqui
	18	2596	40.4	595	22	AAM01318	Alpha prepro-P501S
	19	1677.5	26.1	530	23	AAM50662	Thioredoxin-ubiqui
	20	1522.5	23.7	326	23	ABB77571	Human mast cell re
	21	1420.5	22.1	371	22	AAU69875	Human prostate cDN
	22	1420.5	22.1	371	22	AAM01230	P553S splice varia
	23	1420.5	22.1	371	23	ABB95335	Human P553S splice
	24	1419	22.1	371	22	AAE01362	Human gene 11 enco
	25	1419	22.1	371	23	ABG64105	Human albumin fusi
	26	1403.5	21.9	400	22	AAU69907	Human prostate pro
	27	1403.5	21.9	400	22	AAM01262	Ra12-P501S-E2 cons
	28	1403.5	21.9	400	23	ABB95367	Ra12-P501S-E2 cons
	29	1287	20.1	255	20	AAW85068	Protein encoded by
	30	1287	20.1	255	21	AAB29268	Human prostate-rel
	31	1287	20.1	255	22	AAU04205	Prostate-specific
	32	1217	19.0	252	22	AAE01423	Human secreted pro
	33	1150	17.9	231	21	AAY54369	Amino acid sequenc
c	34	704	11.3	162	21	AAY54370	Amino acid sequenc
	35	673	10.5	129	22	AAE01396	Human gene 11 enco
	36	673	10.5	129	23	ABG64102	Human albumin fusi
	37	519.5	8.1	599	22	ABB60709	Drosophila melanog
	38	466	7.3	748	22	AAM40227	Human polypeptide
c	39	439	7.1	91	22	AAU30352	Novel human secret
	40	426	6.6	123	22	AAU69873	Human prostate cDN
	41	426	6.6	123	22	AAM01228	P553S splice varia
	42	426	6.6	123	23	ABB95333	Human P553S splice
c	43	395	6.3	762	16	AAR80327	Protein polymeric
c	44	395	6.3	762	19	AAW49715	Protein polymer ad
c	45	395	6.3	762	19	AAW57668	Collagen-like poly

ALIGNMENTS

```
RESULT 1
AAW71869
ID   AAW71869 standard; Protein; 553 AA.
XX
AC   AAW71869;
XX
DT   06-JAN-1999  (first entry)
XX
DE   Amino acid encoded by prostate tumour clone L1-12.
XX
KW   Prostate; cancer; tumour; vaccine; immunogen; clone.
XX
OS   Homo sapiens.
XX
PN   WO9837093-A2.
XX
PD   27-AUG-1998.
XX
PF   25-FEB-1998;   98WO-US03492.
```

```
XX
PR   09-FEB-1998;   98US-0020956.
PR   25-FEB-1997;   97US-0806099.
PR   01-AUG-1997;   97US-0904804.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Dillon DC,  Xu J;
XX
DR   WPI; 1998-609886/51.
DR   N-PSDB; AAV61201.
XX
PT   Polypeptides comprising immunogenic portions of prostate proteins -
PT   used in a vaccine for the treatment of prostate cancer
XX
PS   Example 1; Page 82-84; 130pp; English.
XX
CC   The present sequence is an immunogenic portion of a prostate tumour
CC   protein. The immunogen, or the DNA encoding it, can be used as a
CC   vaccine for the treatment of prostate cancer. The immunogen was
CC   isolated from a prostate tumour cDNA library obtained by subtracting
CC   a prostate tumour cDNA expression library with a normal tissue cDNA
CC   library.
XX
SQ   Sequence   553 AA;

Alignment Scores:
Pred. No.:              2.89e-208     Length:         553
Score:                  2861.00       Matches:        553
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            44.58%        Indels:         0
DB:                     19            Gaps:           0

US-09-759-143-110 (1-3410) x AAW71869 (1-553)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||
```

```
Db        541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 2
AAW69385
ID   AAW69385 standard; Protein; 553 AA.
XX
AC   AAW69385;
XX
DT   08-DEC-1998  (first entry)
XX
DE   Prostate tumour specific gene clone L1-12 protein.
XX
KW   Prostate tumour specific gene; human; prostate cancer; detection;
KW   therapy.
XX
OS   Homo sapiens.
XX
PN   WO9837418-A2.
XX
PD   27-AUG-1998.
XX
PF   25-FEB-1998;   98WO-US03690.
XX
PR   09-FEB-1998;   98US-0904809.
PR   25-FEB-1997;   97US-0806596.
PR   01-AUG-1997;   97US-0904809.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Dillon DC,  Xu J;
XX
DR   WPI; 1998-480805/41.
DR   N-PSDB; AAV58586.
XX
PT   Novel human prostate specific tumour protein and fragments - useful
PT   for detecting and treating prostate cancers
XX
PS   Example 1; Page 87-89; 141pp; English.
XX
CC   This sequence is encoded by a human prostate tumour specific gene, and
CC   can be used in the method of the invention. The method is for detecting
CC   prostate cancer comprises contacting a biological sample with an agent
CC   able to bind an immunogenic portion of a prostate protein (such as
CC   this protein sequence). An antibody which binds to an immunogenic
CC   portion of the prostate protein, and the method can be used to detect,
CC   monitor progression of, or treat prostate cancers. The antibody may
CC   also be conjugated to a therapeutic agent for use in therapy of prostate
CC   cancers.
XX
SQ   Sequence   553 AA;

Alignment Scores:
Pred. No.:              2.89e-208       Length:         553
Score:                  2861.00         Matches:        553
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            44.58%          Indels:         0
DB:                     19              Gaps:           0

US-09-759-143-110 (1-3410) x AAW69385 (1-553)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
```

```
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy     1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy     1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy     1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy     1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy     1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy     1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
            |||||||||||||||||||||||||||||||||||||||
Db      541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553
```

RESULT 3
AAB28527

```
ID   AAB28527 standard; Protein; 553 AA.
XX
AC   AAB28527;
XX
DT   07-FEB-2001  (first entry)
XX
DE   Protein encoded by human breast tumour cDNA clone P501S.
XX
KW   Human; breast tumour antigen; cytostatic; immunotherapy;
KW   breast cancer; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200061756-A2.
XX
PD   19-OCT-2000.
XX
PF   10-APR-2000; 2000WO-US09688.
XX
PR   09-APR-1999;   99US-0288950.
PR   02-JUL-1999;   99US-0346327.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Reed SG,  Xu J,  Dillon DC;
XX
DR   WPI; 2000-638568/61.
DR   N-PSDB; AAC79473.
XX
PT   A novel isolated polypeptide comprising an immunogenic portion of a
PT   breast cancer protein useful in the detection and treatment of breast
PT   cancer -
XX
PS   Claim 2; Page 92-93; 95pp; English.
XX
CC   The present sequence is encoded by a cDNA sequence which was isolated
CC   from a breast tumour cDNA library. It is provided in a specification
CC   relating to compounds for immunotherapy and diagnosis of breast cancer.
CC   Breast tumour antigens and the polynucleotides that encode them may be
CC   used in the production of a pharmaceutical composition to be used in the
CC   treatment of breast cancer. Proliferated T cells and incubated antigen
CC   presenting cells are also required. The polypeptides and polynucleotides
CC   may also be used to produce a vaccine.
XX
SQ   Sequence   553 AA;
```

```
Alignment Scores:
Pred. No.:                2.89e-208      Length:          553
Score:                    2861.00        Matches:         553
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              44.58%         Indels:          0
DB:                       21             Gaps:            0
```

US-09-759-143-110 (1-3410) x AAB28527 (1-553)

```
Qy      284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy      344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy      404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy      464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy      524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy      584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy      644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy      704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy      764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy      824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy      884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy      944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy     1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy     1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy     1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy     1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||
Db        541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 4
AAY82002
ID   AAY82002 standard; Protein; 553 AA.
XX
AC   AAY82002;
XX
DT   13-JUN-2000  (first entry)
XX
DE   Human immunogenic prostate tumour protein sequence SEQ ID NO:113.
XX
KW   Human; prostate cancer; diagnosis; tumour; gene therapy; detection;
KW   immunogenic; cytostatic; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200004149-A2.
XX
PD   27-JAN-2000.
XX
PF   14-JUL-1999;   99WO-US15838.
XX
PR   14-JUL-1998;   98US-0115453.
PR   14-JUL-1998;   98US-0116134.
PR   23-SEP-1998;   98US-0159812.
PR   23-SEP-1998;   98US-0159822.
PR   15-JAN-1999;   99US-0232149.
PR   15-JAN-1999;   99US-0232880.
PR   09-APR-1999;   99US-0288946.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Dillon DC,  Harlocker SL,  Yuqiu J,  Xu J,  Mitcham JL;
XX
DR   WPI; 2000-171268/15.
XX
PT   New polypeptide useful for treating and diagnosing prostate cancer
PT   comprises an immunogenic portion of prostate tumor protein -
XX
PS   Claim 3; Page 138-139; 263pp; English.
XX
CC   The present invention describes isolated polypeptides, comprising an
CC   immunogenic portion of a prostate tumour protein (PTP). The polypeptides
CC   and polynucleotides encoding them have cytostatic activity and can be
CC   used in vaccines and in gene therapy. The polypeptides and
CC   polynucleotides encoding them, antigen presenting cells which express
CC   the polypeptides, antibodies against the polypeptides and vaccines
CC   comprising them can be used for inhibiting the development of prostate
CC   cancer in a patient. The polypeptides can be used to generate antibodies
CC   or anti-idiotypic antibodies for passive immuno therapy. A portion of
CC   the polynucleotides encoding the polypeptides can be used as a probe or
CC   to modulate the expression of the polypeptides. AAA06241 to AAA06691 and
CC   AAY82000 to AAY82020 represent sequences used in the exemplification of
CC   the present invention.
XX
SQ   Sequence   553 AA;

Alignment Scores:
Pred. No.:              2.89e-208      Length:         553
Score:                  2861.00        Matches:        553
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            44.58%         Indels:         0
DB:                     21             Gaps:           0

US-09-759-143-110 (1-3410) x AAY82002 (1-553)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||
Db        541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553
```

```
RESULT 5
AAU69763
ID   AAU69763 standard; Protein; 553 AA.
XX
AC   AAU69763;
XX
DT   30-JAN-2002  (first entry)
XX
DE   Human prostate cDNA encoded protein #3.
XX
KW   Human; prostate cancer; cytostatic; immunostimulant; tumour; immunogen.
XX
OS   Homo sapiens.
XX
PN   WO200173032-A2.
XX
PD   04-OCT-2001.
XX
PF   27-MAR-2001; 2001WO-US09919.
XX
PR   27-MAR-2000; 2000US-0536857.
PR   09-MAY-2000; 2000US-0568100.
PR   12-MAY-2000; 2000US-0570737.
PR   13-JUN-2000; 2000US-0593793.
PR   27-JUN-2000; 2000US-0605783.
PR   10-AUG-2000; 2000US-0636215.
PR   29-AUG-2000; 2000US-0651236.
PR   06-SEP-2000; 2000US-0657279.
PR   02-OCT-2000; 2000US-0679426.
PR   10-OCT-2000; 2000US-0685166.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Kalos MD;
PI   Fanger GR,  Retter MW,  Stolk JA,  Day CH,  Vedvick TS,  Carter D;
PI   Li SX,  Wang A,  Skeiky YAW,  Hepler WT,  Henderson RA;
XX
DR   WPI; 2001-639232/73.
DR   N-PSDB; AAS63557.
XX
PT   New human prostate-specific polypeptides and polynucleotides useful for
PT   the diagnosis and treatment of cancer, especially prostate cancer -
XX
PS   Claim 2; Page 269-270; 579pp; English.
XX
CC   The invention relates to isolated prostate-specific
CC   polynucleotides, polypeptides, fusion proteins of the polypeptides,
CC   antibodies raised against the polypeptides (or antigenic epitopes
CC   derived from them) and antigen-presenting cells expressing the
CC   polypeptides. The antibodies are useful for detecting the presence of
CC   cancer, especially prostate cancer. The polypeptides, polynucleotides and
CC   the antigen-presenting cells are useful for stimulating and/or expanding
CC   T cells specific for a tumour protein, and for inhibiting the development
CC   of cancer especially prostate cancer. Compositions comprising the
CC   polynucleotide and/or polypeptide are useful for stimulating an immune
CC   response, and for treating cancer. The oligonucleotide is useful for
CC   detecting cancer. The present sequence is a prostate specific
CC   polypeptide of the invention.
XX
SQ   Sequence   553 AA;

Alignment Scores:
Pred. No.:            2.89e-208      Length:        553
```

```
Score:                    2861.00      Matches:        553
Percent Similarity:       100.00%      Conservative:   0
Best Local Similarity:    100.00%      Mismatches:     0
Query Match:              44.58%       Indels:         0
DB:                       22           Gaps:           0

US-09-759-143-110 (1-3410) x AAU69763 (1-553)

Qy    284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy    344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy    404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy    464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy    524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy    584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy    644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy    704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy    764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy    824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy    884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy    944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy   1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy   1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy   1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy   1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy   1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy   1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy   1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy   1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy   1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy   1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy   1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy   1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy   1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy   1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy   1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy   1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
          |||||||||||||||||||||||||||||||||||||||
Db    541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 6
AAU04961
ID   AAU04961 standard; Protein; 553 AA.
XX
AC   AAU04961;
XX
DT   24-OCT-2001  (first entry)
XX
DE   Human prostate tumour protein L1-12.
XX
KW   Human; prostate tumour protein; prostate cancer.
XX
OS   Homo sapiens.
XX
PN   US6262245-B1.
XX
PD   17-JUL-2001.
XX
PF   25-FEB-1998;   98US-0030607.
XX
PR   25-FEB-1997;   97US-0806099.
PR   01-AUG-1997;   97US-0904804.
PR   09-FEB-1998;   98US-0020956.
XX
PA   (CORI-) CORIXA CORP.
XX
```

```
PI   Xu J,  Dillon DC;
XX
DR   WPI; 2001-440862/47.
DR   N-PSDB; AAS10108.
XX
PT   Novel polynucleotide encoding polypeptide comprising a portion of
PT   prostate tumour protein useful for inhibiting development of prostate
PT   cancer or for treating prostate cancer in a patient  -
XX
PS   Example 1; Column 125-127; 105pp; English.
XX
CC   The sequence is a partial prostate tumour protein, encoded by a prostate
CC   tumour specific cDNA. The DNA is useful for inhibiting the development
CC   of prostate cancer or for treating prostate cancer in a patient.
XX
SQ   Sequence   553 AA;

Alignment Scores:
Pred. No.:              2.89e-208    Length:        553
Score:                  2861.00      Matches:       553
Percent Similarity:     100.00%      Conservative:  0
Best Local Similarity:  100.00%      Mismatches:    0
Query Match:            44.58%       Indels:        0
DB:                     22           Gaps:          0

US-09-759-143-110 (1-3410) x AAU04961 (1-553)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||
Db        541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 7
AAM01117
ID   AAM01117 standard; Protein; 553 AA.
XX
AC   AAM01117;
XX
DT   04-OCT-2001  (first entry)
```

```
XX
DE   Human prostate-specific amino acid sequence L1-12.
XX
KW   Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW   cytostatic; gene therapy; metastasis.
XX
OS   Homo sapiens.
XX
PN   WO200151633-A2.
XX
PD   19-JUL-2001.
XX
PF   16-JAN-2001; 2001WO-US01574.
XX
PR   14-JAN-2000; 2000US-0483672.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Reed SG;
PI   Kalos MD,  Fanger GR,  Day CH,  Retter MW,  Stolk JA,  Skeiky YAW;
PI   Wang A,  Meagher MJ;
XX
DR   WPI; 2001-425873/45.
XX
PT   New polynucleotide encoding a prostate-specific protein, for
PT   diagnosing, monitoring and treating prostate cancer in a patient and
PT   for use in vaccines -
XX
PS   Claim 2; Page 267-268; 543pp; English.
XX
CC   The present invention describes polynucleotide sequences (I) which encode
CC   prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC   and can be used in vaccine production and gene therapy. (I), (II),
CC   antibodies to (II), fusion proteins comprising (II), and isolated
CC   T cells prepared using (I) or (II) are used treat cancer in a patient.
CC   (I) and the antibodies are also used in the detection of cancer in a
CC   patient. The cancer that is diagnosed or treated is particularly
CC   prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
CC   (I) can be used for monitoring the progression of cancer in a patient.
CC   (I) and (II) can also be used to improve diagnostic and therapeutic
CC   methods for prostate cancer. They can indicate the level of metastasis
CC   as well as the prostate volume. AAH93357 to AAH93944 and AAM01115 to
CC   AAM01318 represent polynucleotide and amino acid sequences used in the
CC   exemplification of the present invention.
XX
SQ   Sequence   553 AA;

Alignment Scores:
Pred. No.:              2.89e-208      Length:         553
Score:                  2861.00        Matches:        553
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            44.58%         Indels:         0
DB:                     22             Gaps:           0

US-09-759-143-110 (1-3410) x AAM01117 (1-553)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy         1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy         1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy         1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy         1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy         1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
                |||||||||||||||||||||||||||||||||||||||
Db          541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 8
AAG99002
ID   AAG99002 standard; Protein; 553 AA.
XX
AC   AAG99002;
XX
DT   25-SEP-2001  (first entry)
XX
DE   Human prostate-specific amino acid sequence L1-12/P501S.
XX
KW   Human; prostate cancer; therapy; diagnosis; cat eye syndrome;
KW   chromosome 22q11.2; prostate-specific protein; chromosome 1;
KW   prostate specific antigen; PSA.
XX
OS   Homo sapiens.
XX
PN   WO200134802-A2.
XX
PD   17-MAY-2001.
XX
PF   09-NOV-2000; 2000WO-US30904.
XX
PR   12-NOV-1999;   99US-0439313.
PR   18-NOV-1999;   99US-0443686.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Reed SG;
PI   Kalos MD,  Retter MW,  Stolk JA,  Day CH,  Skeiky YAW,  Wang A;
XX
DR   WPI; 2001-308785/32.
XX
PT   Isolated polypeptide comprising at least an immunogenic portion of a
PT   prostate-specific protein, useful in the diagnosis and therapy of
PT   prostate cancer -
XX
PS   Claim 3; Page 167-168; 325pp; English.
XX
CC   The present invention describes an isolated polypeptide (P1) comprising
CC   at least an immunogenic portion of a prostate-specific protein, or its
CC   variant. Also described are polynucleotides (N1) encoding (P1). (P1) and
CC   (N1) have cytostatic activity and can be used in vaccine production.
CC   The polypeptides, nucleic acids and antibodies from the present
CC   invention are useful in the diagnosis and therapy of prostate cancer.
CC   Prostate specific genes P704P, P712P, P774P, P775P and B305D are located
CC   in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome
CC   region. Prostate specific antigen (PSA) P501S was located on
CC   chromosome 1. AAH84671 to AAH85143 and AAG99000 to AAG99077 represent
CC   polynucleotide and polypeptide sequences used in the exemplification
CC   of the present invention.
XX
SQ   Sequence   553 AA;

Alignment Scores:
Pred. No.:               2.89e-208      Length:         553
Score:                   2861.00        Matches:        553
Percent Similarity:      100.00%        Conservative:   0
Best Local Similarity:   100.00%        Mismatches:     0
Query Match:             44.58%         Indels:         0
DB:                      22             Gaps:           0

US-09-759-143-110 (1-3410) x AAG99002 (1-553)

Qy          284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy          344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy          404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy          464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy          524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy          584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy          644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy          704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy          764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy          824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy          884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy          944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy         1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy         1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy         1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300
```

```
Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||
Db        541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553
```

RESULT 9
AAG62150

```
ID   AAG62150 standard; Protein; 553 AA.
XX
AC   AAG62150;
XX
DT   06-JUL-2001  (first entry)
XX
DE   Human P501S inventive antigen SEQ ID NO: 333.
XX
KW   Human; mouse; immunotherapy; cancer; leukaemia; WT1; Wilm's tumour gene;
KW   chromosome 11p13; zinc finger transcription factor.
XX
OS   Homo sapiens.
XX
PN   WO200125273-A2.
XX
PD   12-APR-2001.
XX
PF   04-OCT-2000; 2000WO-US27465.
XX
PR   04-OCT-1999;   99US-0157459.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky YAW,  Xu J,  Cheever MA,  Reed SG;
XX
DR   WPI; 2001-328324/34.
XX
PT   Polypeptide comprising part of the Wilm's Tumour gene product sequence is
PT   used in the diagnosis and treatment of malignant diseases e.g. leukemia
PT   and cancer associated with WTI  -
XX
PS   Disclosure; Page 212-213; 228pp; English.
XX
CC   The present invention describes compositions comprising peptides derived
CC   from the Wilm's tumour protein WT1 and methods for their use in treating
CC   malignant diseases. Peptides derived from both the murine and human WT1
CC   proteins are provided. The human WT1 gene is found on chromosome 11p13,
CC   and the protein was shown to be a zinc finger transcription factor. The
CC   immunogenic peptides of the invention are particularly useful in the
CC   diagnosis and treatment of cancer and leukaemia. The present sequence is
CC   a polypeptide described in the exemplification of the invention.
XX
SQ   Sequence   553 AA;

Alignment Scores:
Pred. No.:              2.89e-208      Length:        553
Score:                  2861.00        Matches:       553
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            44.58%         Indels:        0
DB:                     22             Gaps:          0

US-09-759-143-110 (1-3410) x AAG62150 (1-553)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||
Db        541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553
```

```
RESULT 10
AAB74800
ID   AAB74800 standard; Protein; 553 AA.
XX
AC   AAB74800;
XX
DT   14-JUN-2001  (first entry)
XX
DE   Prostate tumour antigen predicted amino acid sequence for L1-12.
XX
KW   Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;
KW   prostate cancer; immunogenic; cytostatic; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200125272-A2.
XX
PD   12-APR-2001.
XX
PF   04-OCT-2000; 2000WO-US27464.
XX
PR   04-OCT-1999;    99US-0157455.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Skeiky YAW,  Reed SG,  Cheever MA;
XX
DR   WPI; 2001-245062/25.
DR   N-PSDB; AAH02530.
XX
PT   Prostate specific protein and its encoding polynucleotide, useful for
PT   the treatment and diagnosis of prostate cancer -
XX
PS   Claim 3; Page 157-158; 276pp; English.
XX
CC   The present invention describes an isolated polypeptide (I) comprising
CC   at least an immunogenic portion of a prostate tumour antigen protein or
CC   its variant. (I) have cytostatic activity and can be used in vaccine
CC   production. (I), prostate tumour antigen polynucleotides, an antigen
CC   presenting cell (APC e.g. a dendritic cell) that expresses (I), and a
CC   pharmaceutical composition containing (I) are useful for inhibiting the
CC   development of cancer in a patient. Antibodies specific for prostate
CC   specific proteins and oligonucleotides that hybridise to a
CC   polynucleotide that encodes a prostate specific protein are useful
CC   for detecting the presence or absence of a cancer or monitoring the
CC   progression the progression of a cancer, especially prostate cancer.
CC   AAH02422 to AAH2872, AAB74798 to AAB74821 and AAB74830 are sequences
CC   used in the exemplification of the present invention.
XX
SQ   Sequence   553 AA;

Alignment Scores:
Pred. No.:               2.89e-208      Length:          553
Score:                   2861.00        Matches:         553
Percent Similarity:      100.00%        Conservative:    0
Best Local Similarity:   100.00%        Mismatches:      0
Query Match:             44.58%         Indels:          0
DB:                      22             Gaps:            0

US-09-759-143-110 (1-3410) x AAB74800 (1-553)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60
```

```
Qy    464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy    524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy    584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy    644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy    704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy    764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy    824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy    884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy    944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy   1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy   1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy   1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy   1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy   1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy   1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy   1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy   1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy   1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy   1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy   1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy   1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy   1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy   1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy   1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy   1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
          |||||||||||||||||||||||||||||||||||||||
Db    541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553
```

RESULT 11
ABB77575
ID   ABB77575 standard; Protein; 553 AA.
XX
AC   ABB77575;
XX
DT   30-AUG-2002 (first entry)
XX
DE   Human mast cell related splice variant protein MC14 SEQ ID NO 13.
XX
KW   Human; mast cell; MC; antiallergic; antiinflammatory; antiasthmatic;
KW   vasotropic; dermatological; allergy; hypersensitivity; rhinitis; asthma.
XX
OS   Homo sapiens.
XX
PN   WO200246389-A2.
XX
PD   13-JUN-2002.
XX
PF   07-DEC-2001; 2001WO-US46180.
XX
PR   08-DEC-2000; 2000US-251835P.
PR   14-MAR-2001; 2001US-275479P.
PR   28-MAR-2001; 2001US-279115P.
PR   02-APR-2001; 2001US-280143P.
XX
PA   (UNIO) UCB SA.
XX
PI   Nocka K, Pirozzi G, Einstein R;
XX
DR   WPI; 2002-508560/54.
DR   N-PSDB; ABN81324.
XX
PT   Novel isolated nucleic acids that are differentially expressed in mast
PT   cells in patients with allergic hypersensitivity, encoding proteins
PT   associated with mast cell regranulation and allergic hypersensitivity
PT   -
XX
PS   Claim 31; Page 117-119; 119pp; English.
XX
CC   The invention relates to isolated nucleic acid (ABN81319-ABN81324),
CC   corresponding to genes differentially expressed in mast cells following
CC   activation or in patients with allergic hypersensitivity disease, (I)
CC   that encodes proteins (ABB77569-ABB77575) (II) or a protein fragment of
CC   (II) if at least 6 amino acids. (II) is useful for identifying binding

```
CC   partners. (I) or (II) is useful for diagnosing or treating a disease
CC   state (e.g. allergic hypersensitivity, seasonal rhinitis, asthma,
CC   urticaria or atopic dermatitis or mastocytosis) in a subject which
CC   involves determining the level of expression of (I) or (II). A computer
CC   system, comprising a database containing information identifying the
CC   expression level in a tissue or at least one mast cell of (I), is useful
CC   for presenting information to identify the relative expression level of
CC   (I). (II) is used as a marker to detect, diagnose or identify an allergic
CC   response in a patient. The protein can also serve as a target that
CC   modulate gene expression or activity and as an antigen to raise
CC   polyclonal or monoclonal antibodies. (II) is useful for identifying
CC   agents that modulate expression of the protein or agents, such as
CC   agonists or antagonists. The agonists or antagonists are useful for
CC   modulating biological activity and function of (II) and thus are useful
CC   for alleviating disease conditions such as allergic hypersensitivity,
CC   seasonal rhinitis, asthma, urticaria, atopic dermatitis or mastocytosis.
XX
SQ   Sequence   553 AA;

Alignment Scores:
Pred. No.:              2.89e-208     Length:         553
Score:                  2861.00       Matches:        553
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            44.58%        Indels:         0
DB:                     23            Gaps:           0

US-09-759-143-110 (1-3410) x ABB77575 (1-553)

Qy    284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy    344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy    404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy    464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy    524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy    584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy    644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy    704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy    764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy    824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy    884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy    944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy   1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy   1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy   1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy   1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy   1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy   1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy   1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy   1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy   1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy   1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy   1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy   1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy   1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy   1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy   1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy   1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
          |||||||||||||||||||||||||||||||||||||||
Db    541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 12
ABG61900
ID   ABG61900 standard; Protein; 553 AA.
XX
AC   ABG61900;
```

```
XX
DT   15-AUG-2002  (first entry)
XX
DE   Prostate cancer-associated protein #101.
XX
KW   Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
XX
OS   Mammalia.
XX
PN   WO200230268-A2.
XX
PD   18-APR-2002.
XX
PF   12-OCT-2001; 2001WO-US32045.
XX
PR   13-OCT-2000; 2000US-0687576.
PR   08-DEC-2000; 2000US-0733288.
PR   08-DEC-2000; 2000US-0733742.
PR   24-JAN-2001; 2001US-263957P.
PR   16-MAR-2001; 2001US-276791P.
PR   16-MAR-2001; 2001US-276888P.
PR   06-APR-2001; 2001US-281922P.
PR   24-APR-2001; 2001US-286214P.
PR   30-APR-2001; 2001US-0847046.
PR   04-MAY-2001; 2001US-288589P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Gish KC,  Mack DH,  Wilson KE,  Afar D,  Hevezi P;
XX
DR   WPI; 2002-471335/50.
DR   N-PSDB; ABK92217.
XX
PT   Detecting a prostate cancer-associated transcript in a cell in a
PT   patient, useful for diagnosing prostate cancer (PC) or screening
PT   modulators of PC, by determining if prostate cancer-associated genes
PT   are expressed in a prostate tissue  -
XX
PS   Claim 27; Page 386; 436pp; English.
XX
CC   The present invention relates to methods of detecting a prostate
CC   cancer-associated transcript in a cell from a patient. The method
CC   comprises contacting a biological sample from the patient with
CC   prostate cancer-associated polynucleotides (designated PC genes) that
CC   selectively hybridise to a sequence that is at least 80% identical
CC   to them. The prostate cancer-associated polynucleotide sequences
CC   are differentially expressed in prostate tumour tissue or in
CC   prostate cancer and are derived from the tissues of various
CC   organisms such as humans or other mammals (e.g. mice, sheep and dogs).
CC   The methods of the invention are useful for diagnosing and treating
CC   prostate cancer in mammals. The prostate cancer-associated genes are
CC   useful for diagnosing or treating prostate cancer, as well as for
CC   identifying modulators of prostate cancer or agents that inhibit
CC   prostate cancer. The nucleic acid sequences are particularly useful
CC   in gene therapy, as a vaccine or in antisense applications.
CC   ABG61800-ABG61944 represent prostate cancer-associated proteins.
XX
SQ   Sequence   553 AA;

Alignment Scores:
Pred. No.:              2.89e-208     Length:          553
Score:                  2861.00       Matches:         553
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            44.58%        Indels:          0
DB:                     23            Gaps:            0

US-09-759-143-110 (1-3410) x ABG61900 (1-553)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGACGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||
Db        541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 13
ABB95222
ID   ABB95222 standard; Protein; 553 AA.
XX
AC   ABB95222;
XX
DT   19-JUL-2002  (first entry)
XX
DE   Human L1-12 protein SEQ ID NO 113.
XX
KW   Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
KW   gene therapy.
XX
OS   Homo sapiens.
XX
PN   US2002022248-A1.
XX
PD   21-FEB-2002.
XX
PF   12-JAN-2001; 2001US-0759143.
XX
PR   25-FEB-1997;    97US-0806099.
PR   01-AUG-1997;    97US-0904804.
PR   09-FEB-1998;    98US-0020956.
PR   25-FEB-1998;    98US-0030607.
PR   14-JUL-1998;    98US-0115453.
PR   23-SEP-1998;    98US-0159812.
PR   15-JAN-1999;    99US-0232149.
PR   09-APR-1999;    99US-0288946.
PR   13-JUL-1999;    99US-0352616.
PR   12-NOV-1999;    99US-0439313.
PR   18-NOV-1999;    99US-0443686.
PR   14-JAN-2000;  2000US-0483672.
PR   27-MAR-2000;  2000US-0536857.
PR   09-MAY-2000;  2000US-0568100.
PR   12-MAY-2000;  2000US-0570737.
PR   13-JUN-2000;  2000US-0593793.
PR   27-JUN-2000;  2000US-0605783.
PR   10-AUG-2000;  2000US-0636215.
PR   29-AUG-2000;  2000US-0651236.
PR   06-SEP-2000;  2000US-0657279.
PR   02-OCT-2000;  2000US-0679426.
PR   10-OCT-2000;  2000US-0685166.
XX
PA   (XUJJ/) XU J.
PA   (DILL/) DILLON D C.
PA   (MITC/) MITCHAM J L.
PA   (HARL/) HARLOCKER S L.
PA   (JIAN/) JIANG Y.
PA   (KALO/) KALOS M D.
PA   (FANG/) FANGER G R.
PA   (RETT/) RETTER M W.
PA   (STOL/) STOLK J A.
PA   (DAYC/) DAY C H.
PA   (VEDV/) VEDVICK T S.
PA   (CART/) CARTER D.
PA   (LISX/) LI S X.
PA   (WANG/) WANG A.
PA   (SKEI/) SKEIKY Y A W.
PA   (HEPL/) HEPLER W T.
PA   (HEND/) HENDERSON R A.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Kalos MD;
PI   Fanger GR,  Retter MW,  Stolk JA,  Day CH,  Vedvick TS,  Carter D;
PI   Li SX,  Wang A,  Skeiky YAW,  Hepler WT,  Henderson RA;
XX
DR   WPI; 2002-255649/30.
XX
PT   New prostate-specific polynucleotides for diagnosing and treating
PT   diseases, in particular prostate cancer, and as markers for the
PT   progression of cancer -
XX
PS   Claim 2; SEQ ID NO 113; 87pp; English.
XX
CC   The present invention provides prostate-specific coding sequences and
CC   their encoded proteins. These can be used in the diagnosis and treatment
CC   of cancers, particularly prostate cancer. The present sequence is a
CC   protein described in the invention.
XX
SQ   Sequence   553 AA;

Alignment Scores:
Pred. No.:              2.89e-208     Length:          553
Score:                  2861.00       Matches:         553
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            44.58%        Indels:          0
DB:                     23            Gaps:            0

US-09-759-143-110 (1-3410) x ABB95222 (1-553)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100
```

```
Qy    584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy    644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy    704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy    764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy    824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy    884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy    944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy   1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy   1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy   1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy   1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy   1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy   1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy   1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy   1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy   1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy   1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy   1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy   1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy   1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy   1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy   1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy   1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
          |||||||||||||||||||||||||||||||||||||||
Db    541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 14
AAU10324
ID   AAU10324 standard; Protein; 553 AA.
XX
AC   AAU10324;
XX
DT   14-FEB-2002  (first entry)
XX
DE   Human PROST 03.
XX
KW   Human; PROST 03; metastasis; prostate cancer; tumour; immune response;
KW   cytostatic; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200181577-A2.
XX
PD   01-NOV-2001.
XX
PF   26-APR-2001; 2001WO-US13323.
XX
PR   27-APR-2000; 2000US-200065P.
PR   20-APR-2001; 2001US-0200065.
XX
PA   (SCHD ) SCHERING AG.
XX
PI   Lau T,  Lin RJ,  Parkes D,  Parry G,  Schneider DW,  Steinbrecher R;
PI   Van Heuit PT,  Wu J;
XX
DR   WPI; 2002-041404/05.
DR   N-PSDB; AAS14962.
XX
PT   Novel PROST 03 polypeptides and polynucleotides useful in research,
PT   diagnosis and therapeutic applications, particularly for use in cancer
PT   therapeutics  -
XX
PS   Claim 14; Fig 2; 77pp; English.
XX
CC   The invention relates to an isolated PROST 03 polypeptide (I) and to
CC   the polynucleotide (II) encoding PROST 03. Fragments of (I) were used
CC   to generate antibodies (III) to PROST 03. (III) is useful for selectively
CC   destroying a cell expressing (I), and for treating a disease-state
CC   associated with expression of PROST 03 in a human patient. (III) is
CC   useful for diagnosing metastasis associated with (I), in a subject.
CC   (I) is also useful for diagnosing and treating diseases of cell
CC   proliferation such as prostate cancers. (I) is also useful for generating
CC   antibodies to PROST 03. (III) is useful in detecting the levels of PROST
CC   03 polypeptides in cells and tissues, and in targeting drugs to primary
CC   and metastatic tumours. (I) is also useful for stimulating immune
CC   response to PROST 03 containing cells. (II) is useful in diagnostic
CC   assays for detecting the levels of polynucleotides encoding PROST 03 in
CC   cells and tissues. (II) is useful as DNA probes, as targets for antisense
CC   and ribozyme therapy, and as templates for the production of antisense
CC   polynucleotides. (I) and (II) are useful in research, biological,
```

```
CC   clinical and therapeutic purposes. The present sequence represents
CC   the amino acid sequence of human PROST 03.
XX
SQ   Sequence   553 AA;

Alignment Scores:
Pred. No.:              2.89e-208       Length:         553
Score:                  2861.00         Matches:        553
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            44.58%          Indels:         0
DB:                     23              Gaps:           0

US-09-759-143-110 (1-3410) x AAU10324 (1-553)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||
Db        541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 15
AAU82643
ID   AAU82643 standard; Peptide; 553 AA.
XX
AC   AAU82643;
XX
DT   23-APR-2002  (first entry)
XX
DE   Human breast tumour polypeptide clone #3.
XX
KW   Human; breast tumour polypeptide; breast cancer; cytostatic;
KW   immunostimulant.
XX
OS   Homo sapiens.
XX
PN   WO200198339-A2.
XX
PD   27-DEC-2001.
XX
```

```
PF   12-JUN-2001; 2001WO-US19032.
XX
PR   22-JUN-2000; 2000US-0602877.
PR   12-OCT-2000; 2000US-0687507.
PR   06-FEB-2001; 2001US-0778381.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Reed SG,  Xu J,  Dillon DC,  Retter MW,  Harlocker SL;
XX
DR   WPI; 2002-147792/19.
DR   N-PSDB; ABK29012.
XX
PT   Polynucleotides encoding breast tumour polypeptides, useful for treating
PT   breast cancer or stimulating an immune response -
XX
PS   Claim 2; Page 144-145; 150pp; English.
XX
CC   The invention relates to polynucleotides encoding breast tumour
CC   polypeptides. The sequences are useful for treating cancer, preferably
CC   breast cancer, in a patient or for stimulating an immune response. The
CC   polynucleotides and polypeptides are also useful in the diagnosis and
CC   monitoring of breast cancer. A method for detecting the presence of a
CC   cancer in a patient, comprises obtaining a biological sample from the
CC   patient, contacting the biological sample with a binding agent that binds
CC   to a breast tumour polypeptide, detecting in the sample an amount of
CC   polypeptide that binds to the binding agent, and comparing the amount of
CC   polypeptide to a predetermined cut-off value, therefore determining the
CC   presence of a cancer in the patient. Sequences AAU82641-AAU82655
CC   represent human breast tumour polypeptides of the invention.
XX
SQ   Sequence   553 AA;

Alignment Scores:
Pred. No.:              2.89e-208      Length:        553
Score:                  2861.00        Matches:       553
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            44.58%         Indels:        0
DB:                     23             Gaps:          0

US-09-759-143-110 (1-3410) x AAU82643 (1-553)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520
```

```
Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||
Db        541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 16
AAB74830
ID   AAB74830 standard; Protein; 1079 AA.
XX
AC   AAB74830;
XX
DT   14-JUN-2001  (first entry)
XX
DE   Prostate tumour antigen amino acid sequence for a fusion protein.
XX
KW   Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;
KW   prostate cancer; immunogenic; cytostatic; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200125272-A2.
XX
PD   12-APR-2001.
XX
PF   04-OCT-2000; 2000WO-US27464.
XX
PR   04-OCT-1999;   99US-0157455.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Skeiky YAW,  Reed SG,  Cheever MA;
XX
DR   WPI; 2001-245062/25.
XX
PT   Prostate specific protein and its encoding polynucleotide, useful for
PT   the treatment and diagnosis of prostate cancer -
XX
PS   Disclosure; Page 272-276; 276pp; English.
XX
CC   The present invention describes an isolated polypeptide (I) comprising
CC   at least an immunogenic portion of a prostate tumour antigen protein or
CC   its variant. (I) have cytostatic activity and can be used in vaccine
CC   production. (I), prostate tumour antigen polynucleotides, an antigen
CC   presenting cell (APC e.g. a dendritic cell) that expresses (I), and a
CC   pharmaceutical composition containing (I) are useful for inhibiting the
CC   development of cancer in a patient. Antibodies specific for prostate
CC   specific proteins and oligonucleotides that hybridise to a
CC   polynucleotide that encodes a prostate specific protein are useful
CC   for detecting the presence or absence of a cancer or monitoring the
CC   progression the progression of a cancer, especially prostate cancer.
CC   AAH02422 to AAH2872, AAB74798 to AAB74821 and AAB74830 are sequences
CC   used in the exemplification of the present invention.
XX
SQ   Sequence   1079 AA;

Alignment Scores:
Pred. No.:              3.57e-208      Length:         1079
Score:                  2861.00        Matches:        553
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            44.58%         Indels:         0
DB:                     22             Gaps:           0

US-09-759-143-110 (1-3410) x AAB74830 (1-1079)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 546

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 566

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 586

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 606

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 626

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 646

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 666

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 686

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 706

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 726

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        727 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 746

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 766

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 786

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 806

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 826

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 846

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        847 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 866

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 886

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 906

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          907 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 926

Qy         1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          927 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 946

Qy         1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          947 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 966

Qy         1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          967 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 986

Qy         1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          987 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 1006

Qy         1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1007 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 1026

Qy         1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1027 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 1046

Qy         1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1047 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 1066

Qy         1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
                |||||||||||||||||||||||||||||||||||||||
Db         1067 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 1079

RESULT 17
AAM50661
ID   AAM50661 standard; Protein; 710 AA.
XX
AC   AAM50661;
XX
DT   08-APR-2002  (first entry)
XX
DE   Thioredoxin-ubiquitin-P501S(aa55-553)-His triple fusion protein.
XX
KW   Thioredoxin; ubiquitin; P501S; prostate; tumour; antigen; cancer;
KW   vaccine; therapy; human.
XX
OS   Chimeric - Homo sapiens.
OS   Chimeric - Escherichia coli.
XX
FH   Key             Location/Qualifiers
FT   Protein         1..109
FT                   /label= Thioredoxin
FT   Peptide         110..125
FT                   /label= Linker
FT   Protein         126..199
FT                   /label= Ubiquitin
FT   Cleavage-site   201..202
FT                   /label= UBP1_cleavage_site
FT   Protein         203..702
FT                   /label= P501S
FT   Peptide         703..710
FT                   /label= Histidine_tail
XX
PN   WO200200892-A1.
XX
PD   03-JAN-2002.
XX
PF   19-JUN-2001; 2001WO-EP06952.
XX
PR   26-JUN-2000; 2000GB-0015619.
PR   30-OCT-2000; 2000GB-0026484.
XX
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Cabezon Silva TEV,  Delisse AEF;
XX
DR   WPI; 2002-147888/19.
DR   N-PSDB; ABA91283.
XX
PT   Novel DNA sequence encoding triple fusion protein comprising ubiquitin
PT   fused between thioredoxin and polypeptide of interest, useful for
PT   producing recombinant polypeptide of interest suitable for medicinal
PT   use -
XX
PS   Example 5; Fig 9b; 87pp; English.
XX
CC   The present sequence is that of a triple fusion protein comprising,
CC   from the N-terminal end, Escherichia coli thioredoxin, human
CC   ubiquitin and amino acids 55-553 of human prostate antigen P501S,
CC   followed by a histidine affinity tail.  The triple fusion protein
CC   was produced in E. coli GI724 host cells transformed by an
CC   expression vector comprising DNA (see ABA91283) encoding the fusion.
CC   This is an example of the production of triple fusion proteins of
CC   the invention that comprise ubiquitin fused between thioredoxin and
CC   a protein of interest, in this case P501S(aa55-553).  A claimed
CC   method of producing a recombinant protein of interest involves:
CC   culturing a host cell (preferably E. coli) under conditions which
CC   allow co-expression of the triple fusion and a ubiquitin-specific
CC   endoprotease (especially UBP1 from Saccharomyces cerevisiae);
CC   and recovering the recombinant protein directly from the bacterial
CC   cells after it has been subjected to the action of the
CC   ubiquitin-specific endoprotease in vivo.  In the present case,
CC   the product is a P501-like protein of 509 amino acids.  The
CC   recombinant protein can used as a vaccine for cancer therapy.
XX
SQ   Sequence   710 AA;

Alignment Scores:
Pred. No.:              1.58e-188     Length:        710
Score:                  2601.00       Matches:       501
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            40.53%        Indels:        0
DB:                     23            Gaps:          0

US-09-759-143-110 (1-3410) x AAM50661 (1-710)

Qy          440 ATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGC 499
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          202 MetThrMetValLeuGlyIleGlyProValLeuGlyLeuValCysValProLeuLeuGly 221

Qy          500 TCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTG 559
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          222 SerAlaSerAspHisTrpArgGlyArgTyrGlyArgArgArgProPheIleTrpAlaLeu 241

Qy          560 TCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTG 619
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          242 SerLeuGlyIleLeuLeuSerLeuPheLeuIleProArgAlaGlyTrpLeuAlaGlyLeu 261

Qy          620 CTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTG 679
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          262 LeuCysProAspProArgProLeuGluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeu 281

Qy          680 GACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGG 739
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          282 AspPheCysGlyGlnValCysPheThrProLeuGluAlaLeuLeuSerAspLeuPheArg 301

Qy          740 GACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGC 799
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          302 AspProAspHisCysArgGlnAlaTyrSerValTyrAlaPheMetIleSerLeuGlyGly 321

Qy          800 TGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTG 859
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          322 CysLeuGlyTyrLeuLeuProAlaIleAspTrpAspThrSerAlaLeuAlaProTyrLeu 341
```

```
Qy        860 GGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCA 919
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 GlyThrGlnGluGluCysLeuPheGlyLeuLeuThrLeuIlePheLeuThrCysValAla 361

Qy        920 GCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTG 979
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 AlaThrLeuLeuValAlaGluGluAlaAlaLeuGlyProThrGluProAlaGluGlyLeu 381

Qy        980 TCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAAC 1039
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        382 SerAlaProSerLeuSerProHisCysCysProCysArgAlaArgLeuAlaPheArgAsn 401

Qy       1040 CTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGC 1099
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 LeuGlyAlaLeuLeuProArgLeuHisGlnLeuCysCysArgMetProArgThrLeuArg 421

Qy       1100 CGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTAC 1159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 ArgLeuPheValAlaGluLeuCysSerTrpMetAlaLeuMetThrPheThrLeuPheTyr 441

Qy       1160 ACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAG 1219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        442 ThrAspPheValGlyGluGlyLeuTyrGlnGlyValProArgAlaGluProGlyThrGlu 461

Qy       1220 GCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGC 1279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 AlaArgArgHisTyrAspGluGlyValArgMetGlySerLeuGlyLeuPheLeuGlnCys 481

Qy       1280 GCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGA 1339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 AlaIleSerLeuValPheSerLeuValMetAspArgLeuValGlnArgPheGlyThrArg 501

Qy       1340 GCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCC 1399
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        502 AlaValTyrLeuAlaSerValAlaAlaPheProValAlaAlaGlyAlaThrCysLeuSer 521

Qy       1400 CACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTG 1459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 HisSerValAlaValValThrAlaSerAlaAlaLeuThrGlyPheThrPheSerAlaLeu 541

Qy       1460 CAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCC 1519
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 GlnIleLeuProTyrThrLeuAlaSerLeuTyrHisArgGluLysGlnValPheLeuPro 561

Qy       1520 AAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTG 1579
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        562 LysTyrArgGlyAspThrGlyGlyAlaSerSerGluAspSerLeuMetThrSerPheLeu 581

Qy       1580 CCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGC 1639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 ProGlyProLysProGlyAlaProPheProAsnGlyHisValGlyAlaGlyGlySerGly 601

Qy       1640 CTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTG 1699
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 LeuLeuProProProProAlaLeuCysGlyAlaSerAlaCysAspValSerValArgVal 621

Qy       1700 GTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTC 1759
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        622 ValValGlyGluProThrGluAlaArgValValProGlyArgGlyIleCysLeuAspLeu 641

Qy       1760 GCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCC 1819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        642 AlaIleLeuAspSerAlaPheLeuLeuSerGlnValAlaProSerLeuPheMetGlySer 661

Qy       1820 ATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTG 1879
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 IleValGlnLeuSerGlnSerValThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeu 681

Qy       1880 GTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCA 1939
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        682 ValAlaIleTyrPheAlaThrGlnValValPheAspLysSerAspLeuAlaLysTyrSer 701

Qy       1940 GCG 1942
              |||
Db        702 Ala 702
```

```
RESULT 18
AAM01318
ID   AAM01318 standard; Protein; 595 AA.
XX
AC   AAM01318;
XX
DT   04-OCT-2001  (first entry)
XX
DE   Alpha prepro-P501S recombinant protein amino acid sequence.
XX
KW   Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW   cytostatic; gene therapy; metastasis.
XX
OS   Homo sapiens.
XX
PN   WO200151633-A2.
XX
PD   19-JUL-2001.
XX
PF   16-JAN-2001; 2001WO-US01574.
XX
PR   14-JAN-2000; 2000US-0483672.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Reed SG;
PI   Kalos MD,  Fanger GR,  Day CH,  Retter MW,  Stolk JA,  Skeiky YAW;
PI   Wang A,  Meagher MJ;
XX
DR   WPI; 2001-425873/45.
XX
PT   New polynucleotide encoding a prostate-specific protein, for
PT   diagnosing, monitoring and treating prostate cancer in a patient and
PT   for use in vaccines -
XX
PS   Example 17; Page 541-543; 543pp; English.
XX
CC   The present invention describes polynucleotide sequences (I) which encode
CC   prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC   and can be used in vaccine production and gene therapy. (I), (II),
CC   antibodies to (II), fusion proteins comprising (II), and isolated
CC   T cells prepared using (I) or (II) are used treat cancer in a patient.
CC   (I) and the antibodies are also used in the detection of cancer in a
CC   patient. The cancer that is diagnosed or treated is particularly
CC   prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
CC   (I) can be used for monitoring the progression of cancer in a patient.
CC   (I) and (II) can also be used to improve diagnostic and therapeutic
CC   methods for prostate cancer. They can indicate the level of metastasis
CC   as well as the prostate volume. AAH93357 to AAH93944 and AAM01115 to
CC   AAM01318 represent polynucleotide and amino acid sequences used in the
CC   exemplification of the present invention.
XX
SQ   Sequence   595 AA;

Alignment Scores:
Pred. No.:              3.57e-188     Length:         595
Score:                  2596.00       Matches:        501
Percent Similarity:     96.77%        Conservative:   9
Best Local Similarity:  95.07%        Mismatches:     9
Query Match:            40.45%        Indels:         8
DB:                     22            Gaps:           1

US-09-759-143-110 (1-3410) x AAM01318 (1-595)

Qy        386 GGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACC 445
              |||:::   ::::::         :::      :::      :::|||:::   ::::::
Db         61 GlyLeuLeuPheIleAsnThrThrIleAlaSerIleAlaAlaLysGluGluGlyValSer 80

Qy        446 ------------------------ATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTC 481
```

```
                  ||||||||||||||||||||||||||||||||||||||||||||||
Db         81 LeuGluLysArgGluAlaGluAlaMetValLeuGlyIleGlyProValLeuGlyLeuVal 100

Qy        482 TGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGG 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 CysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArgTyrGlyArgArgArg 120

Qy        542 CCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCC 601
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPheLeuIleProArgAla 140

Qy        602 GGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATC 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 GlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGluLeuAlaLeuLeuIle 160

Qy        662 CTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTG 721
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 LeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThrProLeuGluAlaLeu 180

Qy        722 CTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTC 781
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyrSerValTyrAlaPhe 200

Qy        782 ATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGT 841
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 MetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIleAspTrpAspThrSer 220

Qy        842 GCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATC 901
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPheGlyLeuLeuThrLeuIle 240

Qy        902 TTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACC 961
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAlaAlaLeuGlyProThr 260

Qy        962 GAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCC 1021
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCysCysProCysArgAla 280

Qy       1022 CGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGC 1081
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 ArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHisGlnLeuCysCysArg 300

Qy       1082 ATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATG 1141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 MetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSerTrpMetAlaLeuMet 320

Qy       1142 ACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGA 1201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeuTyrGlnGlyValProArg 340

Qy       1202 GCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTG 1261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 AlaGluProGlyThrGluAlaArgArgHisTyrAspGluGlyValArgMetGlySerLeu 360

Qy       1262 GGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTG 1321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMetAspArgLeuVal 380

Qy       1322 CAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCC 1381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 GlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaAlaPheProValAlaAla 400

Qy       1382 GGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGG 1441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 GlyAlaThrCysLeuSerHisSerValAlaValValThrAlaSerAlaAlaLeuThrGly 420

Qy       1442 TTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAG 1501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 PheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSerLeuTyrHisArgGlu 440

Qy       1502 AAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGC 1561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 LysGlnValPheLeuProLysTyrArgGlyAspThrGlyGlyAlaSerSerGluAspSer 460

Qy       1562 CTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTG 1621
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 LeuMetThrSerPheLeuProGlyProLysProGlyAlaProPheProAsnGlyHisVal 480

Qy       1622 GGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGT 1681
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeuCysGlyAlaSerAlaCys 500

Qy       1682 GATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGG 1741
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 AspValSerValArgValValValGlyGluProThrGluAlaArgValValProGlyArg 520

Qy       1742 GGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCA 1801
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 GlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeuLeuSerGlnValAlaPro 540

Qy       1802 TCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCT 1861
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SerLeuPheMetGlySerIleValGlnLeuSerGlnSerValThrAlaTyrMetValSer 560

Qy       1862 GCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGC 1921
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        561 AlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGlnValValPheAspLysSer 580

Qy       1922 GACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||
Db        581 AspLeuAlaLysTyrSerAla 587
```

```
RESULT 19
AAM50662
ID   AAM50662 standard; Protein; 530 AA.
XX
AC   AAM50662;
XX
DT   08-APR-2002  (first entry)
XX
DE   Thioredoxin-ubiquitin-P501S(aa1-320)-His triple fusion protein.
XX
KW   Thioredoxin; ubiquitin; P501S; prostate; tumour; antigen; cancer;
KW   vaccine; therapy; human.
XX
OS   Chimeric - Homo sapiens.
OS   Chimeric - Escherichia coli.
XX
FH   Key             Location/Qualifiers
FT   Protein         1..109
FT                   /label= Thioredoxin
FT   Peptide         110..125
FT                   /label= Linker
FT   Protein         126..199
FT                   /label= Ubiquitin
FT   Cleavage-site   201..202
FT                   /label= UBP1_cleavage_site
FT   Misc-difference 202..203
FT                   /note= "translation of the DNA sequence of Figure
FT                          12a (ABA91284) places a Val residue between
FT                          Met-202 and Gln-203"
FT   Protein         203..521
FT                   /label= P501S
FT   Misc-difference 208
FT                   /note= "Val at position 208 is not encoded
FT                          by the DNA sequence of Figure 12a
FT                          (ABA91284)"
FT   Peptide         522..530
FT                   /label= Histidine_tail
XX
PN   WO200200892-A1.
XX
PD   03-JAN-2002.
XX
PF   19-JUN-2001; 2001WO-EP06952.
```

```
XX
PR   26-JUN-2000; 2000GB-0015619.
PR   30-OCT-2000; 2000GB-0026484.
XX
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Cabezon Silva TEV,  Delisse AEF;
XX
DR   WPI; 2002-147888/19.
DR   N-PSDB; ABA91284.
XX
PT   Novel DNA sequence encoding triple fusion protein comprising ubiquitin
PT   fused between thioredoxin and polypeptide of interest, useful for
PT   producing recombinant polypeptide of interest suitable for medicinal
PT   use  -
XX
PS   Example 8; Fig 12a; 87pp; English.
XX
CC   The present sequence is that of a triple fusion protein comprising,
CC   from the N-terminal end, Escherichia coli thioredoxin, human
CC   ubiquitin and amino acids 1-320 of human prostate antigen P501S,
CC   followed by a histidine affinity tail.  The triple fusion protein
CC   was produced in E. coli GI724 host cells transformed by an
CC   expression vector comprising DNA (see ABA91284) encoding the fusion.
CC   This is an example of the production of triple fusion proteins of
CC   the invention that comprise ubiquitin fused between thioredoxin and
CC   a protein of interest, in this case P501S(aa55-553).  A claimed
CC   method of producing a recombinant protein of interest involves:
CC   culturing a host cell (preferably E. coli) under conditions which
CC   allow co-expression of the triple fusion and a ubiquitin-specific
CC   endoprotease (especially UBP1 from Saccharomyces cerevisiae);
CC   and recovering the recombinant protein directly from the bacterial
CC   cells after it has been subjected to the action of the
CC   ubiquitin-specific endoprotease in vivo.  In the present case,
CC   expression was controlled by addition of tryptophan.  The
CC   recombinant protein can used as a vaccine for cancer therapy.
XX
SQ   Sequence   530 AA;

Alignment Scores:
Pred. No.:              1.39e-118      Length:        530
Score:                  1677.50        Matches:       318
Percent Similarity:     99.69%         Conservative:  1
Best Local Similarity:  99.38%         Mismatches:    0
Query Match:            26.14%         Indels:        1
DB:                     23             Gaps:          1

US-09-759-143-110 (1-3410) x AAM50662 (1-530)

Qy        287 GTCCAGAGGCTGTGG---GTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              :::||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        202 MetGlnArgLeuTrpValValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 221

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 241

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 261

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 281

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 301

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 321

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 341

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 361

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 381

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        382 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 401

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 421

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 441

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        442 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 461

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 481

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 501

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        502 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 521

RESULT 20
ABB77571
ID   ABB77571 standard; Protein; 326 AA.
XX
AC   ABB77571;
XX
DT   30-AUG-2002  (first entry)
XX
DE   Human mast cell related protein MC14-2 SEQ ID NO 5.
XX
KW   Human; mast cell; MC; antiallergic; antiinflammatory; antiasthmatic;
KW   vasotropic; dermatological; allergy; hypersensitivity; rhinitis; asthma.
XX
OS   Homo sapiens.
XX
PN   WO200246389-A2.
XX
PD   13-JUN-2002.
XX
PF   07-DEC-2001; 2001WO-US46180.
XX
PR   08-DEC-2000; 2000US-251835P.
PR   14-MAR-2001; 2001US-275479P.
PR   28-MAR-2001; 2001US-279115P.
PR   02-APR-2001; 2001US-280143P.
XX
PA   (UNIO ) UCB SA.
XX
PI   Nocka K,  Pirozzi G,  Einstein R;
XX
DR   WPI; 2002-508560/54.
DR   N-PSDB; ABN81320.
XX
```

```
PT   Novel isolated nucleic acids that are differentially expressed in mast
PT   cells in patients with allergic hypersensitivity, encoding proteins
PT   associated with mast cell regranulation and allergic hypersensitivity
PT   -
XX
PS   Claim 31; Page 103-104; 119pp; English.
XX
CC   The invention relates to isolated nucleic acid (ABN81319-ABN81324),
CC   corresponding to genes differentially expressed in mast cells following
CC   activation or in patients with allergic hypersensitivity disease, (I)
CC   that encodes proteins (ABB77569-ABB77575) (II) or a protein fragment of
CC   (II) if at least 6 amino acids. (II) is useful for identifying binding
CC   partners. (I) or (II) is useful for diagnosing or treating a disease
CC   state (e.g. allergic hypersensitivity, seasonal rhinitis, asthma,
CC   urticaria or atopic dermatitis or mastocytosis) in a subject which
CC   involves determining the level of expression of (I) or (II). A computer
CC   system, comprising a database containing information identifying the
CC   expression level in a tissue or at least one mast cell of (I), is useful
CC   for presenting information to identify the relative expression level of
CC   (I). (II) is used as a marker to detect, diagnose or identify an allergic
CC   response in a patient. The protein can also serve as a target that
CC   modulate gene expression or activity and as an antigen to raise
CC   polyclonal or monoclonal antibodies. (II) is useful for identifying
CC   agents that modulate expression of the protein or agents, such as
CC   agonists or antagonists. The agonists or antagonists are useful for
CC   modulating biological activity and function of (II) and thus are useful
CC   for alleviating disease conditions such as allergic hypersensitivity,
CC   seasonal rhinitis, asthma, urticaria, atopic dermatitis or mastocytosis.
XX
SQ   Sequence   326 AA;

Alignment Scores:
Pred. No.:              6.64e-107     Length:        326
Score:                  1522.50       Matches:       300
Percent Similarity:     58.85%        Conservative:  6
Best Local Similarity:  57.69%        Mismatches:    12
Query Match:            23.72%        Indels:        202
DB:                     23            Gaps:          2

US-09-759-143-110 (1-3410) x ABB77571 (1-326)

Qy        347 AACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCT 406
              :::|||:::              |||    :::    |||:::          :::|||:::
Db          9 SerLeuValProLeuProLeuAlaLeuTyrLeuSerGlnProLeuThrHisThrThrSer 28

Qy        407 CTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCA 466
              ||||||                                          ||||||||||||
Db         29 LeuLeu------------------------------------------AlaGlyIleGlyPro 35

Qy        467 GTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGC 526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         36 ValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArg 55

Qy        527 TATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTT 586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         56 TyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPhe 75

Qy        587 CTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAG 646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 LeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGlu 95

Qy        647 CTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACT 706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 LeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThr 115

Qy        707 CCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTAC 766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        116 ProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyr 135

Qy        767 TCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATT 826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 SerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIle 155

Qy        827 GACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGC 886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 AspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPheGly 175

Qy        887 CTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCA 946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 LeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAla 195

Qy        947 GCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGC 1006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 AlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCys 215

Qy       1007 TGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCAC 1066
              |||||||||
Db        216 CysProCys--------------------------------------------------- 218

Qy       1067 CAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGC 1126

Db        218 ------------------------------------------------------------ 218

Qy       1127 TGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTAC 1186

Db        218 ------------------------------------------------------------ 218

Qy       1187 CAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTT 1246

Db        218 ------------------------------------------------------------ 218

Qy       1247 CGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTC 1306

Db        218 ------------------------------------------------------------ 218

Qy       1307 ATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCT 1366

Db        218 ------------------------------------------------------------ 218

Qy       1367 TTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCA 1426

Db        218 ------------------------------------------------------------ 218

Qy       1427 GCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCC 1486

Db        218 ------------------------------------------------------------ 218

Qy       1487 CTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCT 1546

Db        218 ------------------------------------------------------------ 218

Qy       1547 AGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTC 1606
                                               |||||||||||||||||||||||||||
Db        219 -------------------------------Gln-AlaLeuSerLeuGluLeuProSe 227

Qy       1607 CCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGC 1666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 rLeuMetAspThrTrpValLeuGluAlaValAlaCysSerHisLeuHisProArgSerAl 247

Qy       1667 GGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGG 1726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 aGlyProLeuProValMetSerProTyrValTrpTrpTrpValSerProProArgProGl 267

Qy       1727 GTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTG 1786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 yTrpPheArgAlaGlyAlaSerAlaTrpThrSerProSerTrpIleValProSerCysCy 287

Qy       1787 TCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACT 1846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 sProArgTrpProHisProCysLeuTrpAlaProLeuSerSerSerAlaSerLeuSerLe 307

Qy       1847 GCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGG 1904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 uProIleTrpCysLeuProGlnAlaTrpValTrpSerProPheThrLeuLeuHisArg 326

RESULT 21
```

```
AAU69875
ID   AAU69875 standard; Protein; 371 AA.
XX
AC   AAU69875;
XX
DT   30-JAN-2002  (first entry)
XX
DE   Human prostate cDNA encoded protein #69.
XX
KW   Human; prostate cancer; cytostatic; immunostimulant; tumour; immunogen.
XX
OS   Homo sapiens.
XX
PN   WO200173032-A2.
XX
PD   04-OCT-2001.
XX
PF   27-MAR-2001; 2001WO-US09919.
XX
PR   27-MAR-2000; 2000US-0536857.
PR   09-MAY-2000; 2000US-0568100.
PR   12-MAY-2000; 2000US-0570737.
PR   13-JUN-2000; 2000US-0593793.
PR   27-JUN-2000; 2000US-0605783.
PR   10-AUG-2000; 2000US-0636215.
PR   29-AUG-2000; 2000US-0651236.
PR   06-SEP-2000; 2000US-0657279.
PR   02-OCT-2000; 2000US-0679426.
PR   10-OCT-2000; 2000US-0685166.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Kalos MD;
PI   Fanger GR,  Retter MW,  Stolk JA,  Day CH,  Vedvick TS,  Carter D;
PI   Li SX,  Wang A,  Skeiky YAW,  Hepler WT,  Henderson RA;
XX
DR   WPI; 2001-639232/73.
DR   N-PSDB; AAS64038.
XX
PT   New human prostate-specific polypeptides and polynucleotides useful for
PT   the diagnosis and treatment of cancer, especially prostate cancer -
XX
PS   Claim 2; Page 487-488; 579pp; English.
XX
CC   The invention relates to isolated prostate-specific
CC   polynucleotides, polypeptides, fusion proteins of the polypeptides,
CC   antibodies raised against the polypeptides (or antigenic epitopes
CC   derived from them) and antigen-presenting cells expressing the
CC   polypeptides. The antibodies are useful for detecting the presence of
CC   cancer, especially prostate cancer. The polypeptides, polynucleotides and
CC   the antigen-presenting cells are useful for stimulating and/or expanding
CC   T cells specific for a tumour protein, and for inhibiting the development
CC   of cancer especially prostate cancer. Compositions comprising the
CC   polynucleotide and/or polypeptide are useful for stimulating an immune
CC   response, and for treating cancer. The oligonucleotide is useful for
CC   detecting cancer. The present sequence is a prostate specific
CC   polypeptide of the invention.
XX
SQ   Sequence   371 AA;

Alignment Scores:
Pred. No.:              3.72e-99      Length:         371
Score:                  1420.50       Matches:        275
Percent Similarity:     88.58%        Conservative:   12
Best Local Similarity:  84.88%        Mismatches:     23
Query Match:            22.13%        Indels:         14
DB:                     22            Gaps:           1

US-09-759-143-110 (1-3410) x AAU69875 (1-371)

Qy        347 AACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCT 406
              :::|||:::          |||   :::   |||:::       :::|||:::
Db          9 SerLeuValProLeuProLeuAlaLeuTyrLeuSerGlnProLeuThrHisThrThrSer 28

Qy        407 CTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCA 466
              ||||||                                       |||||||||||
Db         29 LeuLeu---------------------------------------AlaGlyIleGlyPro 35

Qy        467 GTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGC 526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         36 ValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArg 55

Qy        527 TATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTT 586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         56 TyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPhe 75

Qy        587 CTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAG 646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 LeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGlu 95

Qy        647 CTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACT 706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 LeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThr 115

Qy        707 CCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTAC 766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        116 ProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyr 135

Qy        767 TCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATT 826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 SerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIle 155

Qy        827 GACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGC 886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 AspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPheGly 175

Qy        887 CTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCA 946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 LeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAla 195

Qy        947 GCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGC 1006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 AlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCys 215

Qy       1007 TGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCAC 1066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 CysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHis 235

Qy       1067 CAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGC 1126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 GlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSer 255

Qy       1127 TGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTAC 1186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 TrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeuTyr 275

Qy       1187 CAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTT 1246
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 GlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGlyLys 295

Qy       1247 CGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTC 1306
                 ::: ||||||              |||   :::      |||   ::::::|||   ::
Db        296 AlaLeu-AlaAlaSerArgGlyTrpCysGlySerArgProProGluThrThrLeuGlyAl 315

Qy       1307 ATGGACCGGC 1316
               :   :::|||
Db        315 aValSerGly 318

RESULT 22
AAM01230
ID   AAM01230 standard; Protein; 371 AA.
XX
AC   AAM01230;
XX
DT   04-OCT-2001  (first entry)
XX
```

```
DE   P553S splice variant P553S-14 amino acid #2.
XX
KW   Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW   cytostatic; gene therapy; metastasis.
XX
OS   Homo sapiens.
XX
PN   WO200151633-A2.
XX
PD   19-JUL-2001.
XX
PF   16-JAN-2001; 2001WO-US01574.
XX
PR   14-JAN-2000; 2000US-0483672.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Reed SG;
PI   Kalos MD,  Fanger GR,  Day CH,  Retter MW,  Stolk JA,  Skeiky YAW;
PI   Wang A,  Meagher MJ;
XX
DR   WPI; 2001-425873/45.
XX
PT   New polynucleotide encoding a prostate-specific protein, for
PT   diagnosing, monitoring and treating prostate cancer in a patient and
PT   for use in vaccines -
XX
PS   Claim 2; Page 464-466; 543pp; English.
XX
CC   The present invention describes polynucleotide sequences (I) which encode
CC   prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC   and can be used in vaccine production and gene therapy. (I), (II),
CC   antibodies to (II), fusion proteins comprising (II), and isolated
CC   T cells prepared using (I) or (II) are used treat cancer in a patient.
CC   (I) and the antibodies are also used in the detection of cancer in a
CC   patient. The cancer that is diagnosed or treated is particularly
CC   prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
CC   (I) can be used for monitoring the progression of cancer in a patient.
CC   (I) and (II) can also be used to improve diagnostic and therapeutic
CC   methods for prostate cancer. They can indicate the level of metastasis
CC   as well as the prostate volume. AAH93357 to AAH93944 and AAM01115 to
CC   AAM01318 represent polynucleotide and amino acid sequences used in the
CC   exemplification of the present invention.
XX
SQ   Sequence   371 AA;

Alignment Scores:
Pred. No.:               3.72e-99      Length:        371
Score:                   1420.50       Matches:       275
Percent Similarity:      88.58%        Conservative:  12
Best Local Similarity:   84.88%        Mismatches:    23
Query Match:             22.13%        Indels:        14
DB:                      22            Gaps:          1

US-09-759-143-110 (1-3410) x AAM01230 (1-371)

Qy        347 AACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCT 406
                 :::|||:::         |||   :::   |||:::      :::|||:::
Db          9 SerLeuValProLeuProLeuAlaLeuTyrLeuSerGlnProLeuThrHisThrThrSer 28

Qy        407 CTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCA 466
              ||||||                                          ||||||||||||
Db         29 LeuLeu------------------------------------------AlaGlyIleGlyPro 35

Qy        467 GTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGC 526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         36 ValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArg 55

Qy        527 TATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTT 586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         56 TyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPhe 75

Qy        587 CTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAG 646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 LeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGlu 95

Qy        647 CTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACT 706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 LeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThr 115

Qy        707 CCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTAC 766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        116 ProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyr 135

Qy        767 TCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATT 826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 SerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIle 155

Qy        827 GACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGC 886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 AspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPheGly 175

Qy        887 CTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCA 946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 LeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAla 195

Qy        947 GCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGC 1006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 AlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCys 215

Qy       1007 TGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCAC 1066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 CysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHis 235

Qy       1067 CAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGC 1126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 GlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSer 255

Qy       1127 TGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTAC 1186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 TrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeuTyr 275

Qy       1187 CAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTT 1246
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 GlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGlyLys 295

Qy       1247 CGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTC 1306
                 ::: ||||||          |||   :::     |||   ::::::|||   ::
Db        296 AlaLeu-AlaAlaSerArgGlyTrpCysGlySerArgProProGluThrThrLeuGlyAl 315

Qy       1307 ATGGACCGGC 1316
              :   :::|||
Db        315 aValSerGly 318

RESULT 23
ABB95335
ID   ABB95335 standard; Protein; 371 AA.
XX
AC   ABB95335;
XX
DT   19-JUL-2002  (first entry)
XX
DE   Human P553S splice variant encoded protein SEQ ID NO 708.
XX
KW   Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
KW   gene therapy.
XX
OS   Homo sapiens.
XX
PN   US2002022248-A1.
XX
PD   21-FEB-2002.
XX
PF   12-JAN-2001; 2001US-0759143.
XX
PR   25-FEB-1997;   97US-0806099.
```

```
PR   01-AUG-1997;    97US-0904804.
PR   09-FEB-1998;    98US-0020956.
PR   25-FEB-1998;    98US-0030607.
PR   14-JUL-1998;    98US-0115453.
PR   23-SEP-1998;    98US-0159812.
PR   15-JAN-1999;    99US-0232149.
PR   09-APR-1999;    99US-0288946.
PR   13-JUL-1999;    99US-0352616.
PR   12-NOV-1999;    99US-0439313.
PR   18-NOV-1999;    99US-0443686.
PR   14-JAN-2000;  2000US-0483672.
PR   27-MAR-2000;  2000US-0536857.
PR   09-MAY-2000;  2000US-0568100.
PR   12-MAY-2000;  2000US-0570737.
PR   13-JUN-2000;  2000US-0593793.
PR   27-JUN-2000;  2000US-0605783.
PR   10-AUG-2000;  2000US-0636215.
PR   29-AUG-2000;  2000US-0651236.
PR   06-SEP-2000;  2000US-0657279.
PR   02-OCT-2000;  2000US-0679426.
PR   10-OCT-2000;  2000US-0685166.
XX
PA   (XUJJ/) XU J.
PA   (DILL/) DILLON D C.
PA   (MITC/) MITCHAM J L.
PA   (HARL/) HARLOCKER S L.
PA   (JIAN/) JIANG Y.
PA   (KALO/) KALOS M D.
PA   (FANG/) FANGER G R.
PA   (RETT/) RETTER M W.
PA   (STOL/) STOLK J A.
PA   (DAYC/) DAY C H.
PA   (VEDV/) VEDVICK T S.
PA   (CART/) CARTER D.
PA   (LISX/) LI S X.
PA   (WANG/) WANG A.
PA   (SKEI/) SKEIKY Y A W.
PA   (HEPL/) HEPLER W T.
PA   (HEND/) HENDERSON R A.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Kalos MD;
PI   Fanger GR,  Retter MW,  Stolk JA,  Day CH,  Vedvick TS,  Carter D;
PI   Li SX,  Wang A,  Skeiky YAW,  Hepler WT,  Henderson RA;
XX
DR   WPI; 2002-255649/30.
XX
PT   New prostate-specific polynucleotides for diagnosing and treating
PT   diseases, in particular prostate cancer, and as markers for the
PT   progression of cancer  -
XX
PS   Claim 2; SEQ ID NO 708; 87pp; English.
XX
CC   The present invention provides prostate-specific coding sequences and
CC   their encoded proteins. These can be used in the diagnosis and treatment
CC   of cancers, particularly prostate cancer. The present sequence is a
CC   protein described in the invention.
XX
SQ   Sequence   371 AA;

Alignment Scores:
Pred. No.:               3.72e-99      Length:          371
Score:                   1420.50       Matches:         275
Percent Similarity:      88.58%        Conservative:    12
Best Local Similarity:   84.88%        Mismatches:      23
Query Match:             22.13%        Indels:          14
DB:                      23            Gaps:            1

US-09-759-143-110 (1-3410) x ABB95335 (1-371)

Qy        347 AACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCT 406
              :::|||:::         |||   :::   |||:::      :::|||:::
Db          9 SerLeuValProLeuProLeuAlaLeuTyrLeuSerGlnProLeuThrHisThrThrSer 28

Qy        407 CTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCA 466
              ||||||                                           |||||||||||
Db         29 LeuLeu------------------------------------------AlaGlyIleGlyPro 35

Qy        467 GTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGC 526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         36 ValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArg 55

Qy        527 TATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTT 586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         56 TyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPhe 75

Qy        587 CTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAG 646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 LeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGlu 95

Qy        647 CTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACT 706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 LeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThr 115

Qy        707 CCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTAC 766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        116 ProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyr 135

Qy        767 TCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATT 826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 SerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIle 155

Qy        827 GACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGC 886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 AspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPheGly 175

Qy        887 CTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCA 946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 LeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAla 195

Qy        947 GCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGC 1006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 AlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCys 215

Qy       1007 TGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCAC 1066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 CysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHis 235

Qy       1067 CAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGC 1126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 GlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSer 255

Qy       1127 TGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTAC 1186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 TrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeuTyr 275

Qy       1187 CAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTT 1246
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 GlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGlyLys 295

Qy       1247 CGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTC 1306
                 ::: ||||||            |||   :::      |||   ::::::|||   ::
Db        296 AlaLeu-AlaAlaSerArgGlyTrpCysGlySerArgProProGluThrThrLeuGlyAl 315

Qy       1307 ATGGACCGGC 1316
              :   :::|||
Db        315 aValSerGly 318

RESULT 24 .
AAE01362
ID   AAE01362 standard; Protein; 371 AA.
XX
AC   AAE01362;
XX
DT   18-JUL-2001  (first entry)
XX
DE   Human gene 11 encoded secreted protein HWBAR14, SEQ ID NO:84.
```

```
XX
KW   Human; secreted protein; proliferative disorder; cancer; tumour; asthma;
KW   foetal abnormality; developmental abnormality; haematopoietic disorder;
KW   immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW   Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;
KW   psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;
KW   inflammation; neurological disorder; Alzheimer's disease; food additive;
KW   angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;
KW   pregnancy-related disorder; endocrine disorder; infection; wound healing;
KW   cell culture; chemotaxis; vulnerary; binding partner identification;
KW   gene therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..48
FT                   /label= Signal_peptide
FT   Protein         49..371
FT                   /label= Mature_human_secreted_protein
FT   Misc-difference 20
FT                   /label= Unknown
FT                   /note= "Encoded by TNC"
XX
PN   WO200134629-A1.
XX
PD   17-MAY-2001.
XX
PF   08-NOV-2000; 2000WO-US30654.
XX
PR   12-NOV-1999;   99US-0164835.
PR   27-JUL-2000; 2000US-0221142.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Komatsoulis GA,  Wei P,  Fiscella M,  Baker KP;
XX
DR   WPI; 2001-308779/32.
DR   N-PSDB; AAD05230.
XX
PT   New nucleic acid encoding one of 21 human secreted proteins for
PT   diagnosing, preventing, treating or ameliorating medical conditions,
PT   such as autoimmune disease and cancer, and used as a food additive or
PT   preservative -
XX
PS   Claim 11; Page 438-440; 490pp; English.
XX
CC   AAD05220-AAD05282 represent cDNAs corresponding to 21 human secreted
CC   protein genes, and AAE01352-AAE01413 represent the proteins they encode.
CC   AAE01415-AAE01433 represent human secreted protein fragments or variants.
CC   The secreted proteins and their genes are useful for preventing, treating
CC   or ameliorating medical conditions, e.g., by protein or gene therapy.
CC   Pathological conditions can be diagnosed by determining the amount of the
CC   new protein in a sample or by determining the presence of mutations in
CC   the new genes. Specific uses are described for each of the 21 genes,
CC   based on the tissues in which they are most highly expressed, and include
CC   developing products for the diagnosis or treatment of proliferative
CC   disorders, cancer, tumours, foetal and developmental abnormalities,
CC   haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
CC   diseases (e.g., rheumatoid arthritis), inflammation, allergies,
CC   neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
CC   cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
CC   psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
CC   angiogenic disorders, kidney disorders, gastrointestinal disorders,
CC   pregnancy-related disorders, endocrine disorders, and infections. The
CC   proteins can also be used to aid wound healing and epithelial cell
CC   proliferation, to prevent skin aging due to sunburn, to maintain organs
CC   before transplantation, for supporting cell culture of primary tissues,
CC   to regenerate tissues, to identify their cognate ligands or binding
CC   partners, and in chemotaxis, and can be used as a food additive or
CC   preservative to modify storage properties. Antibodies specific for a
CC   protein of the invention can be used in alleviating symptoms associated
CC   with the disorders mentioned above, and in diagnostic immunoassays e.g.,
CC   radioimmunoassay or enzyme linked immunosorbent assay (ELISA).
CC   The present sequence represents a human secreted protein of
CC   the invention.
XX
SQ   Sequence   371 AA;

Alignment Scores:
Pred. No.:               4.83e-99      Length:         371
Score:                   1419.00       Matches:        269
Percent Similarity:      95.21%        Conservative:   9
Best Local Similarity:   92.12%        Mismatches:     13
Query Match:             22.11%        Indels:         1
DB:                      22            Gaps:           0

US-09-759-143-110 (1-3410) x AAE01362 (1-371)

Qy        443 ACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCA 502
              :::::::::   ||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 SerLeuLeuAlaGlyIleGlyProValLeuGlyLeuValCysValProLeuLeuGlySer 47

Qy        503 GCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCC 562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         48 AlaSerAspHisTrpArgGlyArgTyrGlyArgArgArgProPheIleTrpAlaLeuSer 67

Qy        563 TTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 LeuGlyIleLeuLeuSerLeuPheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeu 87

Qy        623 TGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGAC 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 CysProAspProArgProLeuGluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAsp 107

Qy        683 TTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGAC 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 PheCysGlyGlnValCysPheThrProLeuGluAlaLeuLeuSerAspLeuPheArgAsp 127

Qy        743 CCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGC 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 ProAspHisCysArgGlnAlaTyrSerValTyrAlaPheMetIleSerLeuGlyGlyCys 147

Qy        803 CTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGC 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 LeuGlyTyrLeuLeuProAlaIleAspTrpAspThrSerAlaLeuAlaProTyrLeuGly 167

Qy        863 ACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCC 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 ThrGlnGluGluCysLeuPheGlyLeuLeuThrLeuIlePheLeuThrCysValAlaAla 187

Qy        923 ACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCG 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 ThrLeuLeuValAlaGluGluAlaAlaLeuGlyProThrGluProAlaGluGlyLeuSer 207

Qy        983 GCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 AlaProSerLeuSerProHisCysCysProCysArgAlaArgLeuAlaPheArgAsnLeu 227

Qy       1043 GGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 GlyAlaLeuLeuProArgLeuHisGlnLeuCysCysArgMetProArgThrLeuArgArg 247

Qy       1103 CTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACG 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 LeuPheValAlaGluLeuCysSerTrpMetAlaLeuMetThrPheThrLeuPheTyrThr 267

Qy       1163 GATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCC 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 AspPheValGlyGluGlyLeuTyrGlnGlyValProArgAlaGluProGlyThrGluAla 287

Qy       1223 CGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCC 1282
              ||||||||||||||||||||        ::: ||||||          |||   :::
Db        288 ArgArgHisTyrAspGluGlyLysAlaLeu-AlaAlaSerArgGlyTrpCysGlySerAr 307

Qy       1283 ATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGC 1316
                 |||   ::::::|||   :::   :::|||
```

```
Db        307 gProProGluThrThrLeuGlyAlaValSerGly 318

RESULT 25
ABG64105
ID   ABG64105 standard; Protein; 371 AA.
XX
AC   ABG64105;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human albumin fusion protein #780.
XX
KW   Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW   human serum albumin; HSA; cancer; reproductive disorder;
KW   digestive disorder; immune disorder; endocrine disorder;
KW   haematopoietic disorder; neural disorder; connective disorder;
KW   cytostatic; antiinfertility; antiinflammatory; antiulcer;
KW   immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;
KW   neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW   osteopathic; antiarthritic.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200177137-A1.
XX
PD   18-OCT-2001.
XX
PF   12-APR-2001; 2001WO-US11988.
XX
PR   12-APR-2000; 2000US-229358P.
PR   25-APR-2000; 2000US-199384P.
PR   21-DEC-2000; 2000US-256931P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Haseltine WA;
XX
DR   WPI; 2002-010886/01.
XX
PT   New fusion protein for treating disease e.g. diabetes comprises an
PT   albumin fused to a therapeutic protein -
XX
PS   Claim 1; Page 1018-1019; 2102pp; English.
XX
CC   The present invention relates to albumin fusion proteins comprising a
CC   therapeutic protein X and human albumin (HA, also known as human serum
CC   albumin, HSA). The proteins are useful for treating a disease or
CC   disorder that may be modulated by therapeutic protein X. The albumin
CC   extends the shelf-life of protein X, and may increase its biological
CC   in vitro/in vivo activity. The protein is useful for treating and
CC   diagnosing disorders such as cancer, reproductive disorders, digestive
CC   disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
CC   (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
CC   (e.g. diabetes), haematopoietic disorders, neural disorders
CC   (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
CC   encephalomyelitis, meningitis, schizophrenia), and connective disorders
CC   (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin
CC   fusion proteins of the invention.
XX
SQ   Sequence   371 AA;

Alignment Scores:
Pred. No.:               4.83e-99      Length:          371
Score:                   1419.00       Matches:         269
Percent Similarity:      95.21%        Conservative:    9
Best Local Similarity:   92.12%        Mismatches:      13
Query Match:             22.11%        Indels:          1
DB:                      23            Gaps:            0

US-09-759-143-110 (1-3410) x ABG64105 (1-371)

Qy        443 ACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCA 502
              ::::::::    ||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 SerLeuLeuAlaGlyIleGlyProValLeuGlyLeuValCysValProLeuLeuGlySer 47

Qy        503 GCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCC 562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         48 AlaSerAspHisTrpArgGlyArgTyrGlyArgArgArgProPheIleTrpAlaLeuSer 67

Qy        563 TTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 LeuGlyIleLeuLeuSerLeuPheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeu 87

Qy        623 TGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGAC 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 CysProAspProArgProLeuGluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAsp 107

Qy        683 TTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGAC 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 PheCysGlyGlnValCysPheThrProLeuGluAlaLeuLeuSerAspLeuPheArgAsp 127

Qy        743 CCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGC 802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 ProAspHisCysArgGlnAlaTyrSerValTyrAlaPheMetIleSerLeuGlyGlyCys 147

Qy        803 CTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGC 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 LeuGlyTyrLeuLeuProAlaIleAspTrpAspThrSerAlaLeuAlaProTyrLeuGly 167

Qy        863 ACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCC 922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 ThrGlnGluGluCysLeuPheGlyLeuLeuThrLeuIlePheLeuThrCysValAlaAla 187

Qy        923 ACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCG 982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 ThrLeuLeuValAlaGluGluAlaAlaLeuGlyProThrGluProAlaGluGlyLeuSer 207

Qy        983 GCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTG 1042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 AlaProSerLeuSerProHisCysCysProCysArgAlaArgLeuAlaPheArgAsnLeu 227

Qy       1043 GGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGG 1102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 GlyAlaLeuLeuProArgLeuHisGlnLeuCysCysArgMetProArgThrLeuArgArg 247

Qy       1103 CTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTACACG 1162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 LeuPheValAlaGluLeuCysSerTrpMetAlaLeuMetThrPheThrLeuPheTyrThr 267

Qy       1163 GATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCC 1222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 AspPheValGlyGluGlyLeuTyrGlnGlyValProArgAlaGluProGlyThrGluAla 287

Qy       1223 CGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCC 1282
              ||||||||||||||||||||       ::: ||||||          |||    :::
Db        288 ArgArgHisTyrAspGluGlyLysAlaLeu-AlaAlaSerArgGlyTrpCysGlySerAr 307

Qy       1283 ATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGC 1316
                 |||    :::::|||    :::    :::|||
Db        307 gProProGluThrThrLeuGlyAlaValSerGly 318

RESULT 26
AAU69907
ID   AAU69907 standard; Protein; 400 AA.
XX
AC   AAU69907;
XX
DT   30-JAN-2002  (first entry)
XX
DE   Human prostate protein/M. tuberculosis Ra12 fusion protein RA12-P501S-E2.
XX
KW   Human; prostate cancer; cytostatic; immunostimulant; tumour; immunogen;
KW   fusion protein.
XX
OS   Chimeric - Homo sapiens.
```

```
OS   Chimeric - Microbacterium tuberculosis.
XX
PN   WO200173032-A2.
XX
PD   04-OCT-2001.
XX
PF   27-MAR-2001; 2001WO-US09919.
XX
PR   27-MAR-2000; 2000US-0536857.
PR   09-MAY-2000; 2000US-0568100.
PR   12-MAY-2000; 2000US-0570737.
PR   13-JUN-2000; 2000US-0593793.
PR   27-JUN-2000; 2000US-0605783.
PR   10-AUG-2000; 2000US-0636215.
PR   29-AUG-2000; 2000US-0651236.
PR   06-SEP-2000; 2000US-0657279.
PR   02-OCT-2000; 2000US-0679426.
PR   10-OCT-2000; 2000US-0685166.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Kalos MD;
PI   Fanger GR,  Retter MW,  Stolk JA,  Day CH,  Vedvick TS,  Carter D;
PI   Li SX,  Wang A,  Skeiky YAW,  Hepler WT,  Henderson RA;
XX
DR   WPI; 2001-639232/73.
DR   N-PSDB; AAS64153.
XX
PT   New human prostate-specific polypeptides and polynucleotides useful for
PT   the diagnosis and treatment of cancer, especially prostate cancer -
XX
PS   Example 17; Page 543-544; 579pp; English.
XX
CC   The invention relates to isolated prostate-specific
CC   polynucleotides, polypeptides, fusion proteins of the polypeptides,
CC   antibodies raised against the polypeptides (or antigenic epitopes
CC   derived from them) and antigen-presenting cells expressing the
CC   polypeptides. The antibodies are useful for detecting the presence of
CC   cancer, especially prostate cancer. The polypeptides, polynucleotides and
CC   the antigen-presenting cells are useful for stimulating and/or expanding
CC   T cells specific for a tumour protein, and for inhibiting the development
CC   of cancer especially prostate cancer. Compositions comprising the
CC   polynucleotide and/or polypeptide are useful for stimulating an immune
CC   response, and for treating cancer. The oligonucleotide is useful for
CC   detecting cancer. The present sequence is fusion protein comprising a
CC   prostate specific polypeptide of the invention.
XX
SQ   Sequence   400 AA;

Alignment Scores:
Pred. No.:              7.39e-98      Length:        400
Score:                  1403.50       Matches:       267
Percent Similarity:     97.45%        Conservative:  0
Best Local Similarity:  97.45%        Mismatches:    2
Query Match:            21.87%        Indels:        5
DB:                     22            Gaps:          1

US-09-759-143-110 (1-3410) x AAU69907 (1-400)

Qy        371 GTGTGTTTGGCCGCAGGC--------------ATCACCTATGTGCCGCCTCTGCTGCTG 415
              |||   ||||||   |||               |||||||||||||||||||||||||||
Db        127 ValThrLeuAlaGluGlyProProAlaGluPheIleThrTyrValProProLeuLeuLeu 146

Qy        416 GAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGC 475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 GluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGlyProValLeuGly 166

Qy        476 CTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGC 535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 LeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArgTyrGlyArg 186

Qy        536 CGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCA 595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 ArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPheLeuIlePro 206

Qy        596 AGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTG 655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 ArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGluLeuAlaLeu 226

Qy        656 CTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAG 715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 LeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThrProLeuGlu 246

Qy        716 GCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTAT 775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 AlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyrSerValTyr 266

Qy        776 GCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGAC 835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 AlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIleAspTrpAsp 286

Qy        836 ACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACC 895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPheGlyLeuLeuThr 306

Qy        896 CTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGC 955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 LeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAlaAlaLeuGly 326

Qy        956 CCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGC 1015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 ProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCysCysProCys 346

Qy       1016 CGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGC 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 ArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHisGlnLeuCys 366

Qy       1076 TGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCA 1135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 CysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSerTrpMetAla 386

Qy       1136 CTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAG 1177
              ||||||||||||||||||||||||||||||||||||||||||
Db        387 LeuMetThrPheThrLeuPheTyrThrAspPheValGlyGlu 400

RESULT 27
AAM01262
ID   AAM01262 standard; Protein; 400 AA.
XX
AC   AAM01262;
XX
DT   04-OCT-2001  (first entry)
XX
DE   Ral2-P501S-E2 construct amino acid sequence.
XX
KW   Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW   cytostatic; gene therapy; metastasis.
XX
OS   Homo sapiens.
XX
PN   WO200151633-A2.
XX
PD   19-JUL-2001.
XX
PF   16-JAN-2001; 2001WO-US01574.
XX
PR   14-JAN-2000; 2000US-0483672.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Reed SG;
PI   Kalos MD,  Fanger GR,  Day CH,  Retter MW,  Stolk JA,  Skeiky YAW;
PI   Wang A,  Meagher MJ;
XX
DR   WPI; 2001-425873/45.
XX
```

```
PT   New polynucleotide encoding a prostate-specific protein, for
PT   diagnosing, monitoring and treating prostate cancer in a patient and
PT   for use in vaccines -
XX
PS   Claim 8; Page 504-506; 543pp; English.
XX
CC   The present invention describes polynucleotide sequences (I) which encode
CC   prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC   and can be used in vaccine production and gene therapy. (I), (II),
CC   antibodies to (II), fusion proteins comprising (II), and isolated
CC   T cells prepared using (I) or (II) are used treat cancer in a patient.
CC   (I) and the antibodies are also used in the detection of cancer in a
CC   patient. The cancer that is diagnosed or treated is particularly
CC   prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
CC   (I) can be used for monitoring the progression of cancer in a patient.
CC   (I) and (II) can also be used to improve diagnostic and therapeutic
CC   methods for prostate cancer. They can indicate the level of metastasis
CC   as well as the prostate volume. AAH93357 to AAH93944 and AAM01115 to
CC   AAM01318 represent polynucleotide and amino acid sequences used in the
CC   exemplification of the present invention.
XX
SQ   Sequence   400 AA;

Alignment Scores:
Pred. No.:              7.39e-98      Length:         400
Score:                  1403.50       Matches:        267
Percent Similarity:     97.45%        Conservative:   0
Best Local Similarity:  97.45%        Mismatches:     2
Query Match:            21.87%        Indels:         5
DB:                     22            Gaps:           1

US-09-759-143-110 (1-3410) x AAM01262 (1-400)

Qy     371 GTGTGTTTGGCCGCAGGC---------------ATCACCTATGTGCCGCCTCTGCTGCTG 415
            |||   |||||   |||                |||||||||||||||||||||||||
Db     127 ValThrLeuAlaGluGlyProProAlaGluPheIleThrTyrValProProLeuLeuLeu 146

Qy     416 GAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGC 475
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     147 GluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGlyProValLeuGly 166

Qy     476 CTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGC 535
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     167 LeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArgTyrGlyArg 186

Qy     536 CGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCA 595
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     187 ArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPheLeuIlePro 206

Qy     596 AGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTG 655
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     207 ArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGluLeuAlaLeu 226

Qy     656 CTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAG 715
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     227 LeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThrProLeuGlu 246

Qy     716 GCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTAT 775
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     247 AlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyrSerValTyr 266

Qy     776 GCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGAC 835
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     267 AlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIleAspTrpAsp 286

Qy     836 ACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACC 895
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     287 ThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPheGlyLeuLeuThr 306

Qy     896 CTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGC 955
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     307 LeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAlaAlaLeuGly 326

Qy     956 CCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGC 1015
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     327 ProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCysCysProCys 346

Qy    1016 CGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGC 1075
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     347 ArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHisGlnLeuCys 366

Qy    1076 TGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCA 1135
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     367 CysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSerTrpMetAla 386

Qy    1136 CTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAG 1177
           ||||||||||||||||||||||||||||||||||||||||||
Db     387 LeuMetThrPheThrLeuPheTyrThrAspPheValGlyGlu 400

RESULT 28
ABB95367
ID   ABB95367 standard; Protein; 400 AA.
XX
AC   ABB95367;
XX
DT   19-JUL-2002  (first entry)
XX
DE   Ra12-P501S-E2 construct SEQ ID NO 852.
XX
KW   Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
KW   gene therapy.
XX
OS   Chimeric - Mycobacterium tuberculosis.
OS   Chimeric - Homo sapiens.
XX
PN   US2002022248-A1.
XX
PD   21-FEB-2002.
XX
PF   12-JAN-2001; 2001US-0759143.
XX
PR   25-FEB-1997;   97US-0806099.
PR   01-AUG-1997;   97US-0904804.
PR   09-FEB-1998;   98US-0020956.
PR   25-FEB-1998;   98US-0030607.
PR   14-JUL-1998;   98US-0115453.
PR   23-SEP-1998;   98US-0159812.
PR   15-JAN-1999;   99US-0232149.
PR   09-APR-1999;   99US-0288946.
PR   13-JUL-1999;   99US-0352616.
PR   12-NOV-1999;   99US-0439313.
PR   18-NOV-1999;   99US-0443686.
PR   14-JAN-2000; 2000US-0483672.
PR   27-MAR-2000; 2000US-0536857.
PR   09-MAY-2000; 2000US-0568100.
PR   12-MAY-2000; 2000US-0570737.
PR   13-JUN-2000; 2000US-0593793.
PR   27-JUN-2000; 2000US-0605783.
PR   10-AUG-2000; 2000US-0636215.
PR   29-AUG-2000; 2000US-0651236.
PR   06-SEP-2000; 2000US-0657279.
PR   02-OCT-2000; 2000US-0679426.
PR   10-OCT-2000; 2000US-0685166.
XX
PA   (XUJJ/) XU J.
PA   (DILL/) DILLON D C.
PA   (MITC/) MITCHAM J L.
PA   (HARL/) HARLOCKER S L.
PA   (JIAN/) JIANG Y.
PA   (KALO/) KALOS M D.
PA   (FANG/) FANGER G R.
PA   (RETT/) RETTER M W.
PA   (STOL/) STOLK J A.
PA   (DAYC/) DAY C H.
PA   (VEDV/) VEDVICK T S.
PA   (CART/) CARTER D.
```

```
PA   (LISX/) LI S X.
PA   (WANG/) WANG A.
PA   (SKEI/) SKEIKY Y A W.
PA   (HEPL/) HEPLER W T.
PA   (HEND/) HENDERSON R A.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Kalos MD;
PI   Fanger GR,  Retter MW,  Stolk JA,  Day CH,  Vedvick TS,  Carter D;
PI   Li SX,  Wang A,  Skeiky YAW,  Hepler WT,  Henderson RA;
XX
DR   WPI; 2002-255649/30.
XX
PT   New prostate-specific polynucleotides for diagnosing and treating
PT   diseases, in particular prostate cancer, and as markers for the
PT   progression of cancer  -
XX
PS   Example 17; SEQ ID NO 852; 87pp; English.
XX
CC   The present invention provides prostate-specific coding sequences and
CC   their encoded proteins. These can be used in the diagnosis and treatment
CC   of cancers, particularly prostate cancer. The present sequence is a
CC   protein described in the invention.
XX
SQ   Sequence   400 AA;

Alignment Scores:
Pred. No.:              7.39e-98      Length:        400
Score:                  1403.50       Matches:       267
Percent Similarity:     97.45%        Conservative:  0
Best Local Similarity:  97.45%        Mismatches:    2
Query Match:            21.87%        Indels:        5
DB:                     23            Gaps:          1

US-09-759-143-110 (1-3410) x ABB95367 (1-400)

Qy        371 GTGTGTTTGGCCGCAGGC---------------ATCACCTATGTGCCGCCTCTGCTGCTG 415
              |||   ||||||   |||               |||||||||||||||||||||||||||
Db        127 ValThrLeuAlaGluGlyProProAlaGluPheIleThrTyrValProProLeuLeuLeu 146

Qy        416 GAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGC 475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 GluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGlyProValLeuGly 166

Qy        476 CTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGC 535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 LeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArgTyrGlyArg 186

Qy        536 CGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCA 595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 ArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPheLeuIlePro 206

Qy        596 AGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTG 655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 ArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGluLeuAlaLeu 226

Qy        656 CTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAG 715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 LeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThrProLeuGlu 246

Qy        716 GCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTAT 775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 AlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyrSerValTyr 266

Qy        776 GCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGAC 835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 AlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIleAspTrpAsp 286

Qy        836 ACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACC 895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 ThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPheGlyLeuLeuThr 306

Qy        896 CTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGC 955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 LeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAlaAlaLeuGly 326

Qy        956 CCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGC 1015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 ProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCysCysProCys 346

Qy       1016 CGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGC 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 ArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHisGlnLeuCys 366

Qy       1076 TGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCA 1135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 CysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSerTrpMetAla 386

Qy       1136 CTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAG 1177
              ||||||||||||||||||||||||||||||||||||||||||
Db        387 LeuMetThrPheThrLeuPheTyrThrAspPheValGlyGlu 400
```

```
RESULT 29
AAW85068
ID   AAW85068 standard; Protein; 255 AA.
XX
AC   AAW85068;
XX
DT   12-FEB-1999  (first entry)
XX
DE   Protein encoded by the consensus sequence of the PS108 gene.
XX
KW   PS108 gene; prostate disease; benign prostatic hyperplasia; BPH;
KW   prostatitis; prostatic intraepithelial neoplasia; PIN; cancer;
KW   drug screening; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO9850567-A1.
XX
PD   12-NOV-1998.
XX
PF   01-MAY-1998;   98WO-US08930.
XX
PR   02-MAY-1997;   97US-0850713.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Billing-Medel PA,  Cohen M,  Colpitts TL,  Friedman PN;
PI   Gordon J,  Granados EN,  Hodges SC,  Klass MR,  Kratochvil JD;
PI   Roberts-Rapp L,  Russell JC,  Stroupe SD;
XX
DR   WPI; 1999-034731/03.
DR   N-PSDB; AAV71181.
XX
PT   New isolated prostate-specific polynucleotides - used to develop
PT   products for the diagnosis and treatment of prostate diseases, e.g.
PT   benign hyperplasia, prostatic or prostate cancer
XX
PS   Claim 17; Pages 99-100; 122pp; English.
XX
CC   The present sequence is encoded by the consensus sequence for a PS108
CC   gene. The sequence is derived from overlapping clones AAV71166-79. The
CC   clone sequences are PS108 gene-specific. They are used in the method
CC   of the invention. The specification describes a method for detecting the
CC   presence of a target PS108 polynucleotide in a test sample. The method
CC   comprises contacting the test sample with at least 1 PS108-specific
CC   polynucleotide or complement, and detecting the presence of the target
CC   PS108 polynucleotide. The products can be used for detecting,
CC   diagnosing, staging, monitoring, prognosticating, in vivo imaging,
CC   preventing or treating, or determining predisposition to diseases or
CC   conditions of the prostate such as benign prostatic hyperplasia (BPH),
CC   prostatitis, prostatic intraepithelial neoplasia (PIN) and cancer. In
CC   particular the products can be used in drug screening and gene therapy.
XX
SQ   Sequence   255 AA;
```

```
Alignment Scores:
Pred. No.:              4.32e-89     Length:        255
Score:                  1287.00      Matches:       255
Percent Similarity:     100.00%      Conservative:  0
Best Local Similarity:  100.00%      Mismatches:    0
Query Match:            20.05%       Indels:        0
DB:                     20           Gaps:          0

US-09-759-143-110 (1-3410) x AAW85068 (1-255)

Qy       1178 GGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGAT 1237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GlyLeuTyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAsp 20

Qy       1238 GAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTC 1297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GluGlyValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPhe 40

Qy       1298 TCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGT 1357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 SerLeuValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSer 60

Qy       1358 GTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTG 1417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ValAlaAlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValVal 80

Qy       1418 ACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACA 1477
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ThrAlaSerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThr 100

Qy       1478 CTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACT 1537
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 LeuAlaSerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThr 120

Qy       1538 GGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGA 1597
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GlyGlyAlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGly 140

Qy       1598 GCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCC 1657
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 AlaProPheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProPro 160

Qy       1658 GCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACC 1717
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AlaLeuCysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThr 180

Qy       1718 GAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCC 1777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GluAlaArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAla 200

Qy       1778 TTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAG 1837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 PheLeuLeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGln 220

Qy       1838 TCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCT 1897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 SerValThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAla 240

Qy       1898 ACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||||||||
Db        241 ThrGlnValValPheAspLysSerAspLeuAlaLysTyrSerAla 255

RESULT 30
AAB29268
ID   AAB29268 standard; Protein; 255 AA.
XX
AC   AAB29268;
XX
DT   07-FEB-2001  (first entry)
XX
DE   Human prostate-related PS108 partial protein sequence SEQ ID NO: 36.
XX
KW   Human; prostate cancer; PS108; antibody; tumour; metastasis.
XX
OS   Homo sapiens.
XX
PN   US6130043-A.
XX
PD   10-OCT-2000.
XX
PF   01-MAY-1998;   98US-0071710.
XX
PR   02-MAY-1997;   97US-0850713.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Cohen M,  Colpitts TL,  Friedman PN,  Gordon J,  Granados EN;
PI   Billing-Medel PA,  Klass MR,  Roberts-Rapp L,  Stroupe SD,  Yu H;
PI   Kratochvil JD,  Russell JC,  Hodges SC;
XX
DR   WPI; 2000-655655/63.
XX
PT   Methods for detecting target prostate-specific polynucleotides or
PT   diseases of the prostate (e.g. prostate cancer), comprising detecting
PT   the presence of any of PS108 nucleic acid sequences in a test sample -
XX
PS   Example 1; Column 85-88; 55pp; English.
XX
CC   The present invention is related to a number of partial coding and
CC   protein sequences for the human prostate tissue protein PS108. These
CC   sequences can be used in the diagnosis and prognosis of prostate
CC   diseases, particularly prostate cancer. They can also be used to produce
CC   antibodies which can be used in treatment. The present sequence is one
CC   of the PS108 partial protein sequences.
XX
SQ   Sequence   255 AA;

Alignment Scores:
Pred. No.:              4.32e-89     Length:        255
Score:                  1287.00      Matches:       255
Percent Similarity:     100.00%      Conservative:  0
Best Local Similarity:  100.00%      Mismatches:    0
Query Match:            20.05%       Indels:        0
DB:                     21           Gaps:          0

US-09-759-143-110 (1-3410) x AAB29268 (1-255)

Qy       1178 GGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGAT 1237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GlyLeuTyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAsp 20

Qy       1238 GAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTC 1297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GluGlyValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPhe 40

Qy       1298 TCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGT 1357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 SerLeuValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSer 60

Qy       1358 GTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTG 1417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ValAlaAlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValVal 80

Qy       1418 ACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACA 1477
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ThrAlaSerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThr 100

Qy       1478 CTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACT 1537
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 LeuAlaSerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThr 120

Qy       1538 GGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGA 1597
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GlyGlyAlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGly 140

Qy       1598 GCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCC 1657
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          141 AlaProPheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProPro 160

Qy         1658 GCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACC 1717
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          161 AlaLeuCysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThr 180

Qy         1718 GAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCC 1777
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          181 GluAlaArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAla 200

Qy         1778 TTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAG 1837
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          201 PheLeuLeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGln 220

Qy         1838 TCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCT 1897
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          221 SerValThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAla 240

Qy         1898 ACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
                |||||||||||||||||||||||||||||||||||||||||||||
Db          241 ThrGlnValValPheAspLysSerAspLeuAlaLysTyrSerAla 255
```

```
RESULT 31
AAU04205
ID   AAU04205 standard; Protein; 255 AA.
XX
AC   AAU04205;
XX
DT   23-OCT-2001  (first entry)
XX
DE   Prostate-specific PS108 polypeptide #1.
XX
KW   Prostate; PS108; immunogen; drug screening; image localisation;
KW   diagnostic; therapeutic; prostate tissue disease; cancer; metastasis;
KW   expressed sequence tag; EST.
XX
OS   Homo sapiens.
XX
PN   US6252047-B1.
XX
PD   26-JUN-2001.
XX
PF   15-MAR-2000; 2000US-0525397.
XX
PR   01-MAY-1998;   98US-0071710.
PR   02-MAY-1997;   97US-0850713.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Billing-Medel PA,  Cohen M,  Colpitts TL,  Friedman PN,  Gordon J;
PI   Granados EN,  Hodges SC,  Klass MR,  Kratochvil JD,  Roberts-Rapp L;
PI   Russell JC,  Stroupe SD,  Yu H;
XX
DR   WPI; 2001-424488/45.
DR   N-PSDB; AAS07155.
XX
PT   Novel PS108 polypeptide useful in assays for detecting antibodies to
PT   prostate tissue, and as immunogens to produce PS108 antibodies  -
XX
PS   Claim 1; Column 85-87; 55pp; English.
XX
CC   The sequence represents the amino acid sequence of prostate-specific
CC   PS108 #1, encoded by PS108 expressed sequence tag (EST) consensus
CC   sequence. The sequence was produced from overlapping PS108 ESTs sequences
CC   to produce a full length consensus sequence. This sequence was then used
CC   to produce the PS108 polypeptide which is useful in assays for detecting
CC   antibodies to prostate tissue, and as an immunogen to produce antibodies.
CC   The polypeptide is useful for screening compounds which specifically bind
CC   to the polypeptide and for screening for drugs, compounds, or any other
CC   agent which can be used to treat diseases associated with PS108. The
CC   antibody is useful to detect, or for image localisation of, PS108 antigen
CC   in a patient, for detecting or diagnosing a disease or condition, as
CC   delivery agents for therapeutic agents as well as for diagnostic tests
CC   and for screening for diseases or conditions associated with PS108,
CC   especially cancer. The antibody is also useful for generating chimeric
CC   antibodies for therapeutic use, for inhibiting the biological activity of
CC   PS108, in therapy (for e.g. to treat prostate tissue disease including
CC   prostate cancer and its metastases), and to detect the presence of any
CC   polypeptide in a test sample which shares one or more antigenic
CC   determinants with the PS108 polypeptide.
XX
SQ   Sequence   255 AA;

Alignment Scores:
Pred. No.:                4.32e-89        Length:          255
Score:                    1287.00         Matches:         255
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              20.05%          Indels:          0
DB:                       22              Gaps:            0

US-09-759-143-110 (1-3410) x AAU04205 (1-255)

Qy         1178 GGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGAT 1237
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            1 GlyLeuTyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAsp 20

Qy         1238 GAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTC 1297
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           21 GluGlyValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPhe 40

Qy         1298 TCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGT 1357
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           41 SerLeuValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSer 60

Qy         1358 GTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTG 1417
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           61 ValAlaAlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValVal 80

Qy         1418 ACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACA 1477
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           81 ThrAlaSerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThr 100

Qy         1478 CTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACT 1537
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          101 LeuAlaSerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThr 120

Qy         1538 GGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGA 1597
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          121 GlyGlyAlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGly 140

Qy         1598 GCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCC 1657
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          141 AlaProPheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProPro 160

Qy         1658 GCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACC 1717
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          161 AlaLeuCysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThr 180

Qy         1718 GAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCC 1777
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          181 GluAlaArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAla 200

Qy         1778 TTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAG 1837
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          201 PheLeuLeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGln 220

Qy         1838 TCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCT 1897
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          221 SerValThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAla 240

Qy         1898 ACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
                |||||||||||||||||||||||||||||||||||||||||||||
Db          241 ThrGlnValValPheAspLysSerAspLeuAlaLysTyrSerAla 255
```

```
RESULT 32
AAE01423
ID   AAE01423 standard; Protein; 252 AA.
XX
AC   AAE01423;
XX
DT   18-JUL-2001  (first entry)
XX
DE   Human secreted protein variant, SEQ ID NO:147.
XX
KW   Human; secreted protein; proliferative disorder; cancer; tumour; asthma;
KW   foetal abnormality; developmental abnormality; haematopoietic disorder;
KW   immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW   Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;
KW   psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;
KW   inflammation; neurological disorder; Alzheimer's disease; food additive;
KW   angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;
KW   pregnancy-related disorder; endocrine disorder; infection; wound healing;
KW   cell culture; chemotaxis; vulnerary; binding partner identification;
KW   gene therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 162
FT                   /note= "Corresponds to any of the naturally occuring
FT                   L-amino acids"
FT   Misc-difference 222
FT                   /note= "Corresponds to any of the naturally occuring
FT                   L-amino acids"
FT   Misc-difference 236
FT                   /note= "Corresponds to any of the naturally occuring
FT                   L-amino acids"
FT   Misc-difference 240
FT                   /note= "Corresponds to any of the naturally occuring
FT                   L-amino acids"
XX
PN   WO200134629-A1.
XX
PD   17-MAY-2001.
XX
PF   08-NOV-2000; 2000WO-US30654.
XX
PR   12-NOV-1999;   99US-0164835.
PR   27-JUL-2000; 2000US-0221142.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Komatsoulis GA,  Wei P,  Fiscella M,  Baker KP;
XX
DR   WPI; 2001-308779/32.
XX
PT   New nucleic acid encoding one of 21 human secreted proteins for
PT   diagnosing, preventing, treating or ameliorating medical conditions,
PT   such as autoimmune disease and cancer, and used as a food additive or
PT   preservative -
XX
PS   Disclosure; Page 476-477; 490pp; English.
XX
CC   AAD05220-AAD05282 represent cDNAs corresponding to 21 human secreted
CC   protein genes, and AAE01352-AAE01413 represent the proteins they encode.
CC   AAE01415-AAE01433 represent human secreted protein fragments or variants.
CC   The secreted proteins and their genes are useful for preventing, treating
CC   or ameliorating medical conditions, e.g., by protein or gene therapy.
CC   Pathological conditions can be diagnosed by determining the amount of the
CC   new protein in a sample or by determining the presence of mutations in
CC   the new genes. Specific uses are described for each of the 21 genes,
CC   based on the tissues in which they are most highly expressed, and include
CC   developing products for the diagnosis or treatment of proliferative
CC   disorders, cancer, tumours, foetal and developmental abnormalities,
CC   haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
CC   diseases (e.g., rheumatoid arthritis), inflammation, allergies,
CC   neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
CC   cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
CC   psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
CC   angiogenic disorders, kidney disorders, gastrointestinal disorders,
CC   pregnancy-related disorders, endocrine disorders, and infections. The
CC   proteins can also be used to aid wound healing and epithelial cell
CC   proliferation, to prevent skin aging due to sunburn, to maintain organs
CC   before transplantation, for supporting cell culture of primary tissues,
CC   to regenerate tissues, to identify their cognate ligands or binding
CC   partners, and in chemotaxis, and can be used as a food additive or
CC   preservative to modify storage properties. Antibodies specific for a
CC   protein of the invention can be used in alleviating symptoms associated
CC   with the disorders mentioned above, and in diagnostic immunoassays e.g.,
CC   radioimmunoassay or enzyme linked immunosorbent assay (ELISA).
CC   The present sequence represents a human secreted protein variant
CC   referred to in the disclosure of the invention.
XX
SQ   Sequence   252 AA;

Alignment Scores:
Pred. No.:              8.68e-84      Length:         252
Score:                  1217.00       Matches:        247
Percent Similarity:     97.24%        Conservative:   0
Best Local Similarity:  97.24%        Mismatches:     5
Query Match:            18.96%        Indels:         2
DB:                     22            Gaps:           0

US-09-759-143-110 (1-3410) x AAE01423 (1-252)

Qy       1250 ATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATG 1309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMet 20

Qy       1310 GACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTC 1369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaAlaPhe 40

Qy       1370 CCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCC 1429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAlaSerAla 60

Qy       1430 GCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTC 1489
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSerLeu 80

Qy       1490 TACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGC 1549
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 TyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGlyAlaSer 100

Qy       1550 AGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCT 1609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 SerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaProPhePro 120

Qy       1610 AATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGG 1669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeuCysGly 140

Qy       1670 GCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTG 1729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 AlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAlaArgVal 160

Qy       1730 GTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCC 1789
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 Val***GlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeuLeuSer 180

Qy       1790 CAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCC 1849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerValThrAla 200

Qy       1850 TATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTA 1909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGlnValVal 220

Qy       1910 TTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGG 1969
```

```
                 |||   ||||||||||||||| ||||||||||||||| |||||||||    |||||||||
Db        221 Phe***LysSerAspLeuAla-AsnThrGlnArgArg-AsnPheGln***IleGlyVal* 240

Qy       1970 AGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGT 2007
                 |||||||||||||||||||||||||    ||||||
Db        240 **GlyLeuProHisTrpValProAlaProCysSerCys 252

RESULT 33
AAY54369
ID   AAY54369 standard; Protein; 231 AA.
XX
AC   AAY54369;
XX
DT   06-APR-2000  (first entry)
XX
DE   Amino acid sequence of a prostate cancer-associated protein.
XX
KW   Prostate cancer-associated gene; Incyte clone 1864683; bone cancer;
KW   cell proliferation; cancer; adrenal gland cancer; bladder cancer;
KW   prostate cancer.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..47
FT                   /note= "signal peptide"
FT   Modified-site   100
FT                   /note= "potential casein kinase II phosphorylation site"
FT   Modified-site   142
FT                   /note= "potential casein kinase II phosphorylation site"
FT   Modified-site   147
FT                   /note= "potential protein kinase C phosphorylation site"
FT   Domain          188..209
FT                   /note= "potential transmembrane domain"
XX
PN   WO9967384-A2.
XX
PD   29-DEC-1999.
XX
PF   15-JUN-1999;   99WO-US13524.
XX
PR   22-JUN-1998;   98US-0102615.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Walker MG,  Volkmuth W,  Klingler TM,  Sprinzak EA;
XX
DR   WPI; 2000-126631/11.
DR   N-PSDB; AAZ45677.
XX
PT   Identifying biomolecules for the diagnosis and treatment of diseases
PT   associated with cell-proliferation -
XX
PS   Claim 10; Page 50-51; 52pp; English.
XX
CC   The present sequence is encoded by a prostate cancer-associated gene,
CC   whose cDNA is represented by Incyte clone number 1864683. The cDNA
CC   sequence is used in the method of the invention. The specification
CC   describes a method for identifying biomolecules for the diagnosis or
CC   treatment of diseases associated with cell proliferation. The method
CC   comprises examining polynucleotides, consisting of prostate cancer-
CC   specific genes, and genes of unknown function, expressed in cDNA
CC   libraries. The patterns of both gene sets are compared to identify
CC   genes of unknown function with similar expression patterns to the
CC   prostate cancer-specific genes. The biomolecules identified by the
CC   method form pharmaceutical compositions useful for the diagnosis and
CC   treatment of diseases associated with cell proliferation. Such diseases
CC   include cancer of the adrenal gland, bladder and bone, but especially
CC   prostate cancer. The method may also be applied using other disease-
CC   specific genes. The prostate cancer-specific genes facilitate the
CC   diagnosis and treatment of cell proliferation disorders.
XX
SQ   Sequence   231 AA;

  Alignment Scores:
  Pred. No.:              1.01e-78        Length:         231
  Score:                  1150.00         Matches:        230
  Percent Similarity:     99.57%          Conservative:   0
  Best Local Similarity:  99.57%          Mismatches:     1
  Query Match:            17.92%          Indels:         0
  DB:                     21              Gaps:           0

US-09-759-143-110 (1-3410) x AAY54369 (1-231)

Qy       1250 ATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATG 1309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMet 20

Qy       1310 GACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTC 1369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaAlaPhe 40

Qy       1370 CCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCC 1429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAlaSerAla 60

Qy       1430 GCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTC 1489
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSerLeu 80

Qy       1490 TACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGC 1549
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 TyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGlyAlaSer 100

Qy       1550 AGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCT 1609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 SerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaProPhePro 120

Qy       1610 AATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGG 1669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeuCysGly 140

Qy       1670 GCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTG 1729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 AlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAlaArgVal 160

Qy       1730 GTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCC 1789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 ValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeuLeuSer 180

Qy       1790 CAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCC 1849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerValThrAla 200

Qy       1850 TATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTA 1909
              ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db        201 TyrMetValSerAlaAlaAlaLeuGlyLeuValAlaIleTyrPheAlaThrGlnValVal 220

Qy       1910 TTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||
Db        221 PheAspLysSerAspLeuAlaLysTyrSerAla 231

RESULT 34
AAY54370
ID   AAY54370 standard; Protein; 162 AA.
XX
AC   AAY54370;
XX
DT   06-APR-2000  (first entry)
XX
DE   Amino acid sequence of a prostate cancer-associated protein.
XX
KW   Prostate cancer-associated gene; bone cancer; cell proliferation;
KW   cancer; adrenal gland cancer; bladder cancer; prostate cancer.
XX
OS   Homo sapiens.
```

```
XX
FH   Key             Location/Qualifiers
FT   Modified-site   4
FT                   /note= "potential N-glycosylation site"
FT   Modified-site   25
FT                   /note= "potential protein kinase C phosphorylation site"
FT   Modified-site   34
FT                   /note= "potential protein kinase C phosphorylation site"
FT   Modified-site   44
FT                   /note= "potential protein kinase C phosphorylation site"
FT   Modified-site   48
FT                   /note= "potential cAMP and cGMP dependent protein
FT                          kinase phosphorylation site"
FT   Domain          83..99
FT                   /note= "potential BLOCK signature of Ly-6/u-PAR"
XX
PN   WO9967384-A2.
XX
PD   29-DEC-1999.
XX
PF   15-JUN-1999;   99WO-US13524.
XX
PR   22-JUN-1998;   98US-0102615.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Walker MG,  Volkmuth W,  Klingler TM,  Sprinzak EA;
XX
DR   WPI; 2000-126631/11.
XX
PT   Identifying biomolecules for the diagnosis and treatment of diseases
PT   associated with cell-proliferation -
XX
PS   Claim 10; Page 51-52; 52pp; English.
XX
CC   The present sequence is encoded by a prostate cancer-associated gene.
CC   The prostate cancer-associated gene cDNA sequence is used in the
CC   method of the invention. The specification describes a method for
CC   identifying biomolecules for the diagnosis or treatment of diseases
CC   associated with cell proliferation. The method comprises examining
CC   polynucleotides, consisting of prostate cancer-specific genes, and
CC   genes of unknown function, expressed in cDNA libraries. The patterns
CC   of both gene sets are compared to identify genes of unknown function
CC   with similar expression patterns to the prostate cancer-specific genes.
CC   The biomolecules identified by the method form pharmaceutical
CC   compositions useful for the diagnosis and treatment of diseases
CC   associated with cell proliferation. Such diseases include cancer of
CC   the adrenal gland, bladder and bone, but especially prostate cancer.
CC   The method may also be applied using other disease-specific genes. The
CC   prostate cancer-specific genes facilitate the diagnosis and treatment
CC   of cell proliferation disorders.
XX
SQ   Sequence   162 AA;

Alignment Scores:
Pred. No.:              5.67e-45      Length:         162
Score:                  704.00        Matches:        144
Percent Similarity:     97.35%        Conservative:   3
Best Local Similarity:  95.36%        Mismatches:     1
Query Match:            11.31%        Indels:         3
DB:                     21            Gaps:           0

US-09-759-143-110 (1-3410) x AAY54370 (1-162)

Qy        447 ATGGTCATGAACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATG 388
              |||||||||||||||||||||||| |||:::|||||| |||||||||||||||||||::::
Db          1 MetValMetAsnPheSerSerThr-ProIleProAla-GlnArgArgHisIleGlyGluS 20

Qy        387 CCTGCGGCCAAACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCT 328
              :: ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 er-AlaAlaLysHisThrSerArgProLysValSerArgLeuThrSerLysSerTrpAla 39

Qy        327 TTCCGGTGCCGCAGCAGGCGGCTCACCCACAGCCTCTGGACCATAGTGGGCCAGGCGGGT 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         40 PheArgCysArgSerArgArgLeuThrHisSerLeuTrpThrIleValGlyGlnAlaGly 59

Qy        267 AGGGCTCAGGGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCAGAAG 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         60 ArgAlaGlnGlyAlaValGlnAlaLeuGlnAsnCysPheValSerAlaLeuLeuGlnLys 79

Qy        207 CTGCGGCCTCTCCTCCTTGCTGCCGCCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCT 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 LeuArgProLeuLeuLeuAlaAlaAlaAsnCysLeuGlyIleSerGlnAlaProIleSer 99

Qy        147 GCCAGCCCTTTGGTGCCGGTCCAGCTTCTCAGCCCATGCTCAACACCTGCTGCTGTGGGG 88
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 AlaSerProLeuValProValGlnLeuLeuSerProCysSerThrProAlaAlaValGly 119

Qy         87 CACCTCAGTGGGGACACGTCTCATCACTCAGATCCTGGCCGAGGCGCGCGGCTGTCACCC 28
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 HisLeuSerGlyAspThrSerHisHisSerAspProGlyArgGlyAlaArgLeuSerPro 139

Qy         27 GGAGCCAGCGCGTGCAGGCTGGTTCCC 1
              ||||||||||||   ||||||||||||
Db        140 GlyAlaSerAla***ArgLeuValPro 148

RESULT 35
AAE01396
ID   AAE01396 standard; Protein; 129 AA.
XX
AC   AAE01396;
XX
DT   18-JUL-2001  (first entry)
XX
DE   Human gene 11 encoded secreted protein HWBAR14, SEQ ID NO:118.
XX
KW   Human; secreted protein; proliferative disorder; cancer; tumour; asthma;
KW   foetal abnormality; developmental abnormality; haematopoietic disorder;
KW   immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW   Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;
KW   psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;
KW   inflammation; neurological disorder; Alzheimer's disease; food additive;
KW   angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;
KW   pregnancy-related disorder; endocrine disorder; infection; wound healing;
KW   cell culture; chemotaxis; vulnerary; binding partner identification;
KW   gene therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1
FT                   /label= Signal_peptide
FT   Protein         2..129
FT                   /label= Mature_human_secreted_protein
XX
PN   WO200134629-A1.
XX
PD   17-MAY-2001.
XX
PF   08-NOV-2000; 2000WO-US30654.
XX
PR   12-NOV-1999;   99US-0164835.
PR   27-JUL-2000; 2000US-0221142.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Komatsoulis GA,  Wei P,  Fiscella M,  Baker KP;
XX
DR   WPI; 2001-308779/32.
DR   N-PSDB; AAD05264.
XX
PT   New nucleic acid encoding one of 21 human secreted proteins for
PT   diagnosing, preventing, treating or ameliorating medical conditions,
PT   such as autoimmune disease and cancer, and used as a food additive or
PT   preservative.-
XX
```

```
PS   Claim 11; Page 456-457; 490pp; English.
XX
CC   AAD05220-AAD05282 represent cDNAs corresponding to 21 human secreted
CC   protein genes, and AAE01352-AAE01413 represent the proteins they encode.
CC   AAE01415-AAE01433 represent human secreted protein fragments or variants.
CC   The secreted proteins and their genes are useful for preventing, treating
CC   or ameliorating medical conditions, e.g., by protein or gene therapy.
CC   Pathological conditions can be diagnosed by determining the amount of the
CC   new protein in a sample or by determining the presence of mutations in
CC   the new genes. Specific uses are described for each of the 21 genes,
CC   based on the tissues in which they are most highly expressed, and include
CC   developing products for the diagnosis or treatment of proliferative
CC   disorders, cancer, tumours, foetal and developmental abnormalities,
CC   haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
CC   diseases (e.g., rheumatoid arthritis), inflammation, allergies,
CC   neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
CC   cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
CC   psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
CC   angiogenic disorders, kidney disorders, gastrointestinal disorders,
CC   pregnancy-related disorders, endocrine disorders, and infections. The
CC   proteins can also be used to aid wound healing and epithelial cell
CC   proliferation, to prevent skin aging due to sunburn, to maintain organs
CC   before transplantation, for supporting cell culture of primary tissues,
CC   to regenerate tissues, to identify their cognate ligands or binding
CC   partners, and in chemotaxis, and can be used as a food additive or
CC   preservative to modify storage properties. Antibodies specific for a
CC   protein of the invention can be used in alleviating symptoms associated
CC   with the disorders mentioned above, and in diagnostic immunoassays e.g.,
CC   radioimmunoassay or enzyme linked immunosorbent assay (ELISA).
CC   The present sequence represents a human secreted protein of
CC   the invention.
XX
SQ   Sequence   129 AA;

Alignment Scores:
Pred. No.:              1.18e-42        Length:         129
Score:                  673.00          Matches:        129
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            10.49%          Indels:         0
DB:                     22              Gaps:           0

US-09-759-143-110 (1-3410) x AAE01396 (1-129)

Qy       2854 AGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTT 2913
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ArgLysValGluGlyGlyAlaSerGlyLeuAsnGlyPheProAsnHisProSerSerLeu 20

Qy       2914 GGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATG 2973
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GlyProAlaTrpPheProProLeuProLeuProSerThrLeuSerArgThrGlyLeuMet 40

Qy       2974 AAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACC 3033
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 LysAlaLeuProLysIleSerProThrProAsnPheProLeuProProThrPheProThr 60

Qy       3034 AGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAA 3093
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SerSerThrThrLeuPheGlyAlaThrAlaGlyProGluAlaGlnSerAlaValSerGln 80

Qy       3094 GCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACT 3153
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AlaPheValHisLeuSerProGlnSerIleSerValLeuGlyGluSerHisThrGluThr 100

Qy       3154 CAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCC 3213
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlnGluHisProLeuProGluLeuArgGluValLeuSerLeuArgGlyGlyLeuSerAla 120

Qy       3214 GTTTGCAATAATGTCGTCTTATTTATT 3240
              |||||||||||||||||||||||||||
Db        121 ValCysAsnAsnValValLeuPheIle 129

RESULT 36
ABG64102
ID   ABG64102 standard; Protein; 129 AA.
XX
AC   ABG64102;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human albumin fusion protein #777.
XX
KW   Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW   human serum albumin; HSA; cancer; reproductive disorder;
KW   digestive disorder; immune disorder; endocrine disorder;
KW   haematopoietic disorder; neural disorder; connective disorder;
KW   cytostatic; antiinfertility; antiinflammatory; antiulcer;
KW   immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;
KW   neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW   osteopathic; antiarthritic.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200177137-A1.
XX
PD   18-OCT-2001.
XX
PF   12-APR-2001; 2001WO-US11988.
XX
PR   12-APR-2000; 2000US-229358P.
PR   25-APR-2000; 2000US-199384P.
PR   21-DEC-2000; 2000US-256931P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Haseltine WA;
XX
DR   WPI; 2002-010886/01.
XX
PT   New fusion protein for treating disease e.g. diabetes comprises an
PT   albumin fused to a therapeutic protein -
XX
PS   Claim 1; Page 1016-1017; 2102pp; English.
XX
CC   The present invention relates to albumin fusion proteins comprising a
CC   therapeutic protein X and human albumin (HA, also known as human serum
CC   albumin, HSA). The proteins are useful for treating a disease or
CC   disorder that may be modulated by therapeutic protein X. The albumin
CC   extends the shelf-life of protein X, and may increase its biological
CC   in vitro/in vivo activity. The protein is useful for treating and
CC   diagnosing disorders such as cancer, reproductive disorders, digestive
CC   disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
CC   (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
CC   (e.g. diabetes), haematopoietic disorders, neural disorders
CC   (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
CC   encephalomyelitis, meningitis, schizophrenia), and connective disorders
CC   (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin
CC   fusion proteins of the invention.
XX
SQ   Sequence   129 AA;

Alignment Scores:
Pred. No.:              1.18e-42        Length:         129
Score:                  673.00          Matches:        129
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            10.49%          Indels:         0
DB:                     23              Gaps:           0

US-09-759-143-110 (1-3410) x ABG64102 (1-129)

Qy       2854 AGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTT 2913
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ArgLysValGluGlyGlyAlaSerGlyLeuAsnGlyPheProAsnHisProSerSerLeu 20
```

```
Qy       2914 GGCCCAGCCTGGTTCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATG 2973
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GlyProAlaTrpPheProProLeuProLeuProSerThrLeuSerArgThrGlyLeuMet 40

Qy       2974 AAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACC 3033
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 LysAlaLeuProLysIleSerProThrProAsnPheProLeuProProThrPheProThr 60

Qy       3034 AGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAA 3093
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SerSerThrThrLeuPheGlyAlaThrAlaGlyProGluAlaGlnSerAlaValSerGln 80

Qy       3094 GCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACT 3153
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AlaPheValHisLeuSerProGlnSerIleSerValLeuGlyGluSerHisThrGluThr 100

Qy       3154 CAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCC 3213
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlnGluHisProLeuProGluLeuArgGluValLeuSerLeuArgGlyGlyLeuSerAla 120

Qy       3214 GTTTGCAATAATGTCGTCTTATTTATT 3240
              |||||||||||||||||||||||||||
Db        121 ValCysAsnAsnValValLeuPheIle 129
```

```
RESULT 37
ABB60709
ID   ABB60709 standard; Protein; 599 AA.
XX
AC   ABB60709;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 8919.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL04812.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 8919; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   599 AA;

  Alignment Scores:
  Pred. No.:              8.24e-31      Length:          599
  Score:                  519.50        Matches:         150
  Percent Similarity:     39.43%        Conservative:    100
  Best Local Similarity:  23.66%        Mismatches:      201
  Query Match:            8.09%         Indels:          183
  DB:                     22            Gaps:            14

US-09-759-143-110 (1-3410) x ABB60709 (1-599)

Qy        317 CGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGT 376
               |||      :::   ::::::       ::::::   :::          :::|||
Db         46 ArgLysThrArgPheGluMetPheArgLeuSerAlaIleAlaMetAlaIleGluPheAla 65

Qy        377 TTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAG 436
                 ||||||       ::::::|||   |||:::||||||:::::::||||||:::    |||
Db         66 TyrAlaAlaGluThrSerPheValSerProIleLeuLeuGlnIleGlyValAspHisLys 85

Qy        437 TTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTA 496
                 |||:::|||     |||:::   |||::::::|||        |||||||||
Db         86 HisMetSerMetThrTrpGlyLeuSerProLeuIleGlyPhePheMetSerProLeuLeu 105

Qy        497 GGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCA 556
               ||||||  ||||||        :::   |||:::|||||||||||||||    |||
Db        106 GlySerIleSerAspArgCysLysLeuArgTrpGlyArgArgArgProIleIleSerIle 125

Qy        557 CTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCA-------------------- 595
               ||||||  |||||||:::        |||    |||:::|||
Db        126 LeuSerPheGlyIleMetCysGlyLeuIleLeuValProTyrGlyLysAspLeuGlyLeu 145

Qy        596 ---------------------------------------------------------AGG 598

Db        146 LeuLeuGlyAspAlaGlyTyrThrTyrAlaGluSerAlaLeuAsnPheThrSerSerSer 165

Qy        599 GCCGGCTGGCTAGCAGGGCTGCTGTGC-----------------CCGGATCCCAGGCCC 640
                 |||   :::|||   |||:::                          |||
Db        166 GlyGlySerValAlaAlaLeuValSerGlyGluAlaThrThrGlyProSerAlaSerAsp 185

Qy        641 CTGGAGCTGGCA------CTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 694
                 :::   |||      |||   |||||||||:::   |||||||||||
Db        186 TyrLysPheAlaValIleLeuThrIleLeuGlyMetValLeuLeuAspPheAspAlaAsp 205

Qy        695 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 754
                 |||   ||||||                 |||   |||:::         |||:::
Db        206 ThrCysGlnThrProAlaArgThrTyrLeuLeuAspMetCys---ValProGluGluGln 224

Qy        755 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 814
                 :::|||   :::::::::|||             ||||||   :::||||||
Db        225 ProLysAlaMetThrMetPheAlaLeuPheAlaGlyPheGlyGlyThrIleGlyTyrAla 244

Qy        815 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 874
              :::      :::||||||:::|||:::   :::         ::::::|||
Db        245 IleGlyGlyValAspTrpGluThrThrHisIleGlySerPheMetGlyGlyAsnIlePro 264

Qy        875 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTG------ 928
                 :::|||   |||:::|||:::||||||         |||         |||:::
Db        265 ThrValPheThrLeuValThrIleIlePheAlaValCysTyrLeuIleThrValThrThr 284

Qy        929 --------------------CTGGTGGCTGAGGAGGCAGCGCTGGGCCCC--------- 958
                                  |||:::   ::::::         |||   |||
Db        285 PheArgGluIleProLeuProLeuIleGluGlnAspGluLeuLeuArgProLeuSerGlu 304

Qy        958 ------------------------------------------------------------ 958

Db        305 GlnAlaIleLysLysGluLeuLysLysLysAsnAsnThrIleTyrTyrIleGlnGluThr 324

Qy        959 ----------------------------------------ACCGAGCCAGCAGAAGGG 976
                                                      |||      :::|||
Db        325 ThrGlnLeuGluLeuGlnMetAlaSerAspAspProLysArgLeuGluAlaLeuGlnGly 344

Qy        977 CTGTCGGCCCCCTCCTTGTCGCCC------------------------------------ 1000
```

```
                              ||||||
Db        345 SerTyrGlnAsnGlyTyrSerProAlaValGluLysGlnGlyLysSerGlnAspLeuGlu 364

Qy       1001 ------------CACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCC 1048
                          :::     |||       :::       :::::::::
Db        365 ThrGlnSerAspTyrAspAlaProValSerLeuLysAlaTyrLeuLysSerIlePheIle 384

Qy       1049 CTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTC 1108
                                                 ||||||  ::::::|||  |||
Db        385 --------------------------------MetProTyrSerMetArgMetLeuAla 393

Qy       1109 GTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTC 1168
              :::     |||     ||||||      :::|||:::    |||:::::||||||||||
Db        394 LeuThrAsnLeuPheCysTrpMetGlyHisValThrTyrCysLeuTyrPheThrAspPhe 413

Qy       1169 GTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGA 1228
              |||||||||  ::::::   |||   |||   |||   |||   ::::||||||
Db        414 ValGlyGluAlaValPheHisGlyAspProThrAlaAlaProAsnSerGluAlaAlaLeu 433

Qy       1229 CACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCC 1288
              :::|||:::    |||||||||    |||        |||:::     :::
Db        434 AsnTyrGluAlaGlyValArgPheGlyCysTrpGlyMetAlaIleTyrAlaPheSerCys 453

Qy       1289 CTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTAT 1348
                ::::::||||||   :::    :::|||::::::   |||||||||:::||||||||||
Db        454 SerIleTyrSerLeuSerValThrLysLeuMetLysTrpPheGlyThrLysAlaValTyr 473

Qy       1349 TTGGCCAGTGTGGCAGCTTTCCCTGTG---------GCTGCCGGTGCCACATGCCTGTCC 1399
              ::::::   :::      :::   :::                |||
Db        474 IleSerGlyMetIleTyrTyrGlyIleGlyMetLeuValLeuGlyLeuTrpProThrLys 493

Qy       1400 CACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTG 1459
                 |||   |||   :::   ||||||            |||      :::      :::
Db        494 TrpGlyValLeuValPheSerThrSerAla---------GlyIleLeuTyrGlyThrIle 510

Qy       1460 CAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCC 1519
                 :::|||:::   |||   :::   ||||||      |||      |||
Db        511 PheThrValProPheIleLeuValAlaArgTyrHis---AlaLysAsnCysPheSerIle 529

Qy       1520 AAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTG 1579
              |||
Db        530 Lys--------------------------------------------------------- 530

Qy       1580 CCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGC 1639
                                           ||||||
Db        531 ----------------------------AsnGly-------------------------- 532

Qy       1640 CTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTG 1699

Db        532 ------------------------------------------------------------ 532

Qy       1700 GTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCG---------GGCCGGGGCATCTGC 1750
                                        :::||||||          |||||||:::
Db        533 ------------------------GluIleValProLeuLysGlnAlaArgGlyLeuGly 544

Qy       1751 CTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTT 1810
                |||:::||||||:::   |||       ::::::|||:::       ||||||
Db        545 ThrAspValAlaIleIleSerSerMetValPheIleAlaGlnLeuIleValSerLeuSer 564

Qy       1811 ATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGC 1870
              :::|||   :::|||             :::       |||   :::   ::::::
Db        565 ValGlyProLeuValSerTrpMetAspThrThrCysAlaValLeuTyrAlaSerThrPhe 584

Qy       1871 CTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTT 1912
              |||      :::|||     .|||     |||::::::
Db        585 LeuSerPheLeuAlaAlaIleAlaAlaMetPheValLeuTyr 598
```

RESULT 38

```
AAM40227
ID   AAM40227 standard; Protein; 748 AA.
XX
AC   AAM40227;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 3372.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   N-PSDB; AAI59383.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Example 5; SEQ ID NO 3372; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence   748 AA;

Alignment Scores:
Pred. No.:              1e-26          Length:         748
Score:                  466.00         Matches:        173
Percent Similarity:     34.00%         Conservative:   83
Best Local Similarity:  22.97%         Mismatches:     223
Query Match:            7.26%          Indels:         275
DB:                     22             Gaps:           19

US-09-759-143-110 (1-3410) x AAM40227 (1-748)

Qy        222 CCGAGACGAAGCAGTTCTGGAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCA 281
                |||:::|||                  |||      ||||||   |||      |||   |||
```

```
Db       50 ProLysArgArgLysCysIleArgProSerProProProProProAsnThrProCysPro 69

Qy      282 CTATG----------GTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCC 331
             |||                 |||       |||       |||
Db       70 LeuGluLeuValAspPheGlyAspLeuHisProGlnArgSerPheArg----------- 85

Qy      332 CAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATC 391
            :::||||||             :::   |||||||:::|||            |||
Db       86 GluLeuLeuPheAsnGlyCysIleLeuPheGlyIleGluPheSerTyrAlaMetGluThr 105

Qy      392 ACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTG 451
               |||||   |||:::|||||::::::|||:::   ::::::     ::::::|||
Db      106 AlaTyrValThrProValLeuLeuGlnMetGlyLeuProAspGlnLeuTyrSerLeuVal 125

Qy      452 CTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGAC 511
                  |||   |||:::||||||   :::      ||||||||||||:::   ||||||
Db      126 TrpPheIleSerProIleLeuGlyPheLeuLeuGlnProLeuLeuGlyAlaTrpSerAsp 145

Qy      512 CACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATC 571
                     |||:::||||||||||||||||||||      |||::::::|||
Db      146 ArgCysThrSerArgPheGlyArgArgArgProPheIleLeuValLeuAlaIleGlyAla 165

Qy      572 CTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGAT 631
            ||||||   |||   |||:::         |||     |||:::         |||
Db      166 LeuLeuGlyLeuSerLeuLeuLeuAsn---GlyArgAspIleGlyIleAlaLeuAlaAsp 184

Qy      632 ---------CCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGAC 682
                               |||   |||   :::   ||||||   |||:::|||
Db      185 ValThrGlyAsnHisLysTrpGlyLeuLeuLeuThrValCysGlyValValLeuMetAsp 204

Qy      683 TTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGAC 742
             |||                |||      |||   :::   |||:::
Db      205 PheSerAlaAspSerAlaAspAsnProSerHisAlaTyrMetMetAspValCysSerPro 224

Qy      743 CCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGC 802
              |||      |||         :::::::::|||   :::      |||||||||
Db      225 AlaAspGlnAspArg---GlyLeuAsnIleHisAlaLeuLeuAlaGlyLeuGlyGlyGly 243

Qy      803 CTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGC 862
               ||||||::::::   |||   ||||||   :::               ||||||
Db      244 PheGlyTyrValValGlyGlyIleHisTrpAspLysThrGlyPheGlyArgAlaLeuGly 263

Qy      863 ACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTC------ACCTGCGTA 916
               |||         ::::::   |||   |||   :::   |||      |||   |||
Db      264 GlyGlnLeuArgValIleTyr---LeuPheThrAlaValThrLeuSerValThrThrVal 282

Qy      917 GCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGG 976
                     :::   :::   |||            |||:::|||
Db      283 LeuThrLeuValSerIleProGluArgProLeuArgProProSerGluLysArgAlaAla 302

Qy      977 CTGTCGGCCCCCTCCTTG-----------TCGCCCCACTGCTGTCCA----------- 1012
            :::   :::|||||||||         ||||||      |||
Db      303 MetLysSerProSerLeuProLeuProProSerProProValLeuProGluGluGlyPro 322

Qy     1012 ------------------------------------------------------------ 1012

Db      323 GlyAspSerLeuProSerHisThrAlaThrAsnPheSerSerProIleSerProProSer 342

Qy     1012 ------------------------------------------------------------ 1012

Db      343 ProLeuThrProLysTyrGlySerPheIleSerArgAspSerSerLeuThrGlyIleSer 362

Qy     1013 ------------------------------------------------TGC------ 1015
                                                            |||
Db      363 GluPheAlaSerSerPheGlyThrAlaAsnIleAspSerValLeuIleAspCysPheThr 382

Qy     1015 ------------------------------------------------------------ 1015

Db      383 GlyGlyHisAspSerTyrLeuAlaIleProGlySerValProArgProProIleSerVal 402

Qy     1016 ---------CGGGCCCGCTTGGCTTTC-----------CGGAACCTG----------- 1042
                     ||||||      |||            |||   |||
Db      403 SerPheProArgAlaProAspGlyPheTyrArgGlnAspArgGlyLeuLeuGluGlyArg 422

Qy     1043 ---GGCGCCCTGCTT--------------------------------------------- 1054
               |||||||||
Db      423 GluGlyAlaLeuThrSerGlyCysAspGlyAspIleLeuArgValGlySerLeuAspThr 442

Qy     1055 ------CCCCGG------------------------------------------------ 1060
                  ||||||
Db      443 SerLysProArgSerSerGlyIleLeuLysArgProGlnThrLeuAlaIleProAspAla 462

Qy     1060 ------------------------------------------------------------ 1060

Db      463 AlaGlyGlyGlyGlyProGluThrSerArgArgArgAsnValThrPheSerGlnGlnVal 482

Qy     1060 ------------------------------------------------------------ 1060

Db      483 AlaAsnIleLeuLeuAsnGlyValLysTyrGluSerGluLeuThrGlySerSerGluArg 502

Qy     1061 ------------------CTGCACCAGCTGTGC---------TGCCGCATGCCCCGCACC 1093
                              :::   :::||||||         |||   ||||||:::
Db      503 AlaGluGlnProLeuSerValGlyArgLeuCysSerThrIleCysAsnMetProLysAla 522

Qy     1094 CTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTG 1153
            ||||||   |||   |||                  |||:::::::            |||
Db      523 LeuArgThrLeuCysValAsnHisPheLeuGlyTrpLeuSerPheGluGlyMetLeuLeu 542

Qy     1154 TTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGC 1213
            ||||||||||||||:::||||||   ::::::|||||   |||:::|||
Db      543 PheTyrThrAspPheMetGlyGluValValPheGlnGlyAspProLysAlaProHisThr 562

Qy     1214 ACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTG 1273
            :::||||||   :::   |||:::   ||||||   ||||||      |||:::   :::
Db      563 SerGluAlaTyrGlnLysTyrAsnSerGlyValThrMetGlyCysTrpGlyMetCysIle 582

Qy     1274 CAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGC 1333
                     :::      :::|||   ::::::::::::|||   :::
Db      583 TyrAlaPheSerAlaAlaPheTyrSerAlaIleLeuGluLysLeuGluGluPheLeuSer 602

Qy     1334 ACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGC 1393
               |||   :::|||      :::      ||||||   :::      |||
Db      603 ValArgThrLeuTyrPheIleAlaTyrLeuAlaPheGlyLeuGlyThrGlyLeuAlaThr 622

Qy     1394 CTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCA 1453
            ||||||   ::::::   ||||||   :::                  |||   ||||||
Db      623 LeuSerArgAsnLeuTyrValValLeuSerLeuCysIleThrTyrGlyIleLeuPheSer 642

Qy     1454 GCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTC 1513
               |||      ||||||||||:::|||      |||::::::   |||
Db      643 ThrLeuCysThrLeuProTyrSerLeuLeuCysAspTyrTyrGlnSerLys--------- 659

Qy     1514 CTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGC 1573
                  |||:::   |||               ||||||   |||
Db      660 ------LysPheAlaGly---------------SerSerAlaAspGlyThrArg------ 670

Qy     1574 TTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGC 1633

Db      670 ------------------------------------------------------------ 670

Qy     1634 AGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTA 1693

Db      670 ------------------------------------------------------------ 670

Qy     1694 CGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTG 1753
                                                        |||||||:::   :::
Db      671 ------------------------------------------ArgGlyMetGlyVal 675

Qy     1754 GACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATG 1813
            |||::::::::::|||         :::   |||:::|||:::         ||||||   :::
Db      676 AspIleSerLeuLeuSerCysGlnTyrPheLeuAlaGlnIleLeuValSerLeuValLeu 695

Qy     1814 GGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTG 1873
            |||                                 :::   ||||||   |||
Db      696 GlyPro------------------------------LeuThrSerAlaValGlySer 704
```

```
Qy       1874 GGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTT 1912
                             :::||||||:::::: |||  |||
Db        705 AlaAsnGlyValMetTyrPheSerSerLeuValSerPhe 717

RESULT 39
AAU30352
ID   AAU30352 standard; Protein; 91 AA.
XX
AC   AAU30352;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Novel human secreted protein #843.
XX
KW   Human; vaccination; gene therapy; nutritional supplement;
KW   stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW   immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200179449-A2.
XX
PD   25-OCT-2001.
XX
PF   16-APR-2001; 2001WO-US08656.
XX
PR   18-APR-2000; 2000US-0552929.
PR   26-JAN-2001; 2001US-0770160.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-611725/70.
XX
PT   Nucleic acids encoding a range of human polypeptides, useful in genetic
PT   vaccination, testing and therapy -
XX
PS   Claim 20; Page 282; 765pp; English.
XX
CC   The invention relates to novel human secreted polypeptides. The
CC   polypeptides and antibodies to the polypeptides are useful for
CC   determining the presence of or predisposition to a disease associated
CC   with altered levels of polypeptide. The polypeptides are also useful for
CC   identifying agents (agonists and antagonists) that bind to them. Cells
CC   expressing the proteins are useful for identifying a therapeutic agent
CC   for use in treatment of a pathology related to aberrant expression or
CC   physiological interactions of the polypeptide. Vectors comprising
CC   the nucleic acids encoding the polypeptides and cells genetically
CC   engineered to express them are also useful for producing the proteins.
CC   The proteins are useful in genetic vaccination, testing and
CC   therapy, and can be used as nutritional supplements. They may be used to
CC   increase stem cell proliferation; to regulate haematopoiesis; and in
CC   bone, cartilage, tendon  and/or nerve tissue growth or regeneration;
CC   immune suppression and/or stimulation; as anti-inflammatory agents; and
CC   in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
CC   sequences of novel human secreted proteins of the invention.
XX
SQ   Sequence   91 AA;

Alignment Scores:
Pred. No.:               5.71e-25      Length:         91
Score:                   439.00        Matches:        86
Percent Similarity:      95.60%        Conservative:   1
Best Local Similarity:   94.51%        Mismatches:     4
Query Match:             7.05%         Indels:         1
DB:                      22            Gaps:           0

US-09-759-143-110 (1-3410) x AAU30352 (1-91)

Qy       3106 GATGGACAAAGGCTTGGGAAACCGCACTTTGTGGTTCTGGTCCTGCAGTAGCTCCAAACA 3047
              |||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db          1 AspGlyGlnArgLeuGlyLysProHisPheValLeuLeuValLeuGln***LeuGlnThr 20

Qy       3046 GGGTTGTGGAGCTGGTGGGGAAAGTTGGGGTAGGGGAAAG-TTGGGGGTAGGGGAAATT 2988
              ||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db         21 GlyLeuTrpSerTrpTrpGlyLysLeuGlyValGlyGluMetLeuGlyValGlyGluIle 40

Qy       2987 TTGGGCAGTGCCTTCATCAGCCCAGTCCTAGAGAGAGTAGAGGGGAGTGGAAGTGGGGGG 2928
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 LeuGlySerAlaPheIleSerProValLeuGluArgValGluGlySerGlySerGlyGly 60

Qy       2927 AACCAGGCTGGGCCAAGAGAAGAGGGGTGGTTAGGGAAGCCGTTGAGACCTGAAGCCCCA 2868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AsnGlnAlaGlyProArgGluGluGlyTrpLeuGlyLysProLeuArgProGluAlaPro 80

Qy       2867 CCCTCTACCTTCCTTCAACACCCTAACCTTGGG 2835
              ||||||||||||:::||||||     ||||||
Db         81 ProSerThrPheIleGlnHisHisThrLeuGly 91

RESULT 40
AAU69873
ID   AAU69873 standard; Protein; 123 AA.
XX
AC   AAU69873;
XX
DT   30-JAN-2002  (first entry)
XX
DE   Human prostate cDNA encoded protein #67.
XX
KW   Human; prostate cancer; cytostatic; immunostimulant; tumour; immunogen.
XX
OS   Homo sapiens.
XX
PN   WO200173032-A2.
XX
PD   04-OCT-2001.
XX
PF   27-MAR-2001; 2001WO-US09919.
XX
PR   27-MAR-2000; 2000US-0536857.
PR   09-MAY-2000; 2000US-0568100.
PR   12-MAY-2000; 2000US-0570737.
PR   13-JUN-2000; 2000US-0593793.
PR   27-JUN-2000; 2000US-0605783.
PR   10-AUG-2000; 2000US-0636215.
PR   29-AUG-2000; 2000US-0651236.
PR   06-SEP-2000; 2000US-0657279.
PR   02-OCT-2000; 2000US-0679426.
PR   10-OCT-2000; 2000US-0685166.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Kalos MD;
PI   Fanger GR,  Retter MW,  Stolk JA,  Day CH,  Vedvick TS,  Carter D;
PI   Li SX,  Wang A,  Skeiky YAW,  Hepler WT,  Henderson RA;
XX
DR   WPI; 2001-639232/73.
DR   N-PSDB; AAS64041.
XX
PT   New human prostate-specific polypeptides and polynucleotides useful for
PT   the diagnosis and treatment of cancer, especially prostate cancer -
XX
PS   Claim 2; Page 486-487; 579pp; English.
XX
CC   The invention relates to isolated prostate-specific
CC   polynucleotides, polypeptides, fusion proteins of the polypeptides,
CC   antibodies raised against the polypeptides (or antigenic epitopes
CC   derived from them) and antigen-presenting cells expressing the
CC   polypeptides. The antibodies are useful for detecting the presence of
CC   cancer, especially prostate cancer. The polypeptides, polynucleotides and
CC   the antigen-presenting cells are useful for stimulating and/or expanding
CC   T cells specific for a tumour protein, and for inhibiting the development
CC   of cancer especially prostate cancer. Compositions comprising the
CC   polynucleotide and/or polypeptide are useful for stimulating an immune
CC   response, and for treating cancer. The oligonucleotide is useful for
```

```
CC   detecting cancer. The present sequence is a prostate specific
CC   polypeptide of the invention.
XX
SQ   Sequence   123 AA;

Alignment Scores:
Pred. No.:               6.07e-24      Length:         123
Score:                   426.00        Matches:        87
Percent Similarity:      97.83%        Conservative:   3
Best Local Similarity:   94.57%        Mismatches:     2
Query Match:             6.64%         Indels:         0
DB:                      22            Gaps:           0

US-09-759-143-110 (1-3410) x AAU69873 (1-123)

Qy       1250 ATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATG 1309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMet 20

Qy       1310 GACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTC 1369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaAlaPhe 40

Qy       1370 CCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCC 1429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAlaSerAla 60

Qy       1430 GCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTC 1489
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSerLeu 80

Qy       1490 TACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATAC 1525
              |||||||||||||||||||||   :::   ::::::
Db         81 TyrHisArgGluLysGlnValLeuIleGlyGlnTrp 92

RESULT 41
AAM01228
ID   AAM01228 standard; Protein; 123 AA.
XX
AC   AAM01228;
XX
DT   04-OCT-2001  (first entry)
XX
DE   P553S splice variant amino acid P553S-6.
XX
KW   Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW   cytostatic; gene therapy; metastasis.
XX
OS   Homo sapiens.
XX
PN   WO200151633-A2.
XX
PD   19-JUL-2001.
XX
PF   16-JAN-2001; 2001WO-US01574.
XX
PR   14-JAN-2000; 2000US-0483672.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Reed SG;
PI   Kalos MD,  Fanger GR,  Day CH,  Retter MW,  Stolk JA,  Skeiky YAW;
PI   Wang A,  Meagher MJ;
XX
DR   WPI; 2001-425873/45.
XX
PT   New polynucleotide encoding a prostate-specific protein, for
PT   diagnosing, monitoring and treating prostate cancer in a patient and
PT   for use in vaccines -
XX
PS   Claim 2; Page 463-464; 543pp; English.
XX
CC   The present invention describes polynucleotide sequences (I) which encode
CC   prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC   and can be used in vaccine production and gene therapy. (I), (II),
CC   antibodies to (II), fusion proteins comprising (II), and isolated
CC   T cells prepared using (I) or (II) are used treat cancer in a patient.
CC   (I) and the antibodies are also used in the detection of cancer in a
CC   patient. The cancer that is diagnosed or treated is particularly
CC   prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
CC   (I) can be used for monitoring the progression of cancer in a patient.
CC   (I) and (II) can also be used to improve diagnostic and therapeutic
CC   methods for prostate cancer. They can indicate the level of metastasis
CC   as well as the prostate volume. AAH93357 to AAH93944 and AAM01115 to
CC   AAM01318 represent polynucleotide and amino acid sequences used in the
CC   exemplification of the present invention.
XX
SQ   Sequence   123 AA;

Alignment Scores:
Pred. No.:               6.07e-24      Length:         123
Score:                   426.00        Matches:        87
Percent Similarity:      97.83%        Conservative:   3
Best Local Similarity:   94.57%        Mismatches:     2
Query Match:             6.64%         Indels:         0
DB:                      22            Gaps:           0

US-09-759-143-110 (1-3410) x AAM01228 (1-123)

Qy       1250 ATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATG 1309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMet 20

Qy       1310 GACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTC 1369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaAlaPhe 40

Qy       1370 CCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCC 1429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAlaSerAla 60

Qy       1430 GCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTC 1489
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSerLeu 80

Qy       1490 TACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATAC 1525
              |||||||||||||||||||||   :::   ::::::
Db         81 TyrHisArgGluLysGlnValLeuIleGlyGlnTrp 92

RESULT 42
ABB95333
ID   ABB95333 standard; Protein; 123 AA.
XX
AC   ABB95333;
XX
DT   19-JUL-2002  (first entry)
XX
DE   Human P553S splice variant encoded protein SEQ ID NO 706.
XX
KW   Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
KW   gene therapy.
XX
OS   Homo sapiens.
XX
PN   US2002022248-A1.
XX
PD   21-FEB-2002.
XX
PF   12-JAN-2001; 2001US-0759143.
XX
PR   25-FEB-1997;   97US-0806099.
PR   01-AUG-1997;   97US-0904804.
PR   09-FEB-1998;   98US-0020956.
PR   25-FEB-1998;   98US-0030607.
PR   14-JUL-1998;   98US-0115453.
PR   23-SEP-1998;   98US-0159812.
```

```
PR   15-JAN-1999;   99US-0232149.
PR   09-APR-1999;   99US-0288946.
PR   13-JUL-1999;   99US-0352616.
PR   12-NOV-1999;   99US-0439313.
PR   18-NOV-1999;   99US-0443686.
PR   14-JAN-2000; 2000US-0483672.
PR   27-MAR-2000; 2000US-0536857.
PR   09-MAY-2000; 2000US-0568100.
PR   12-MAY-2000; 2000US-0570737.
PR   13-JUN-2000; 2000US-0593793.
PR   27-JUN-2000; 2000US-0605783.
PR   10-AUG-2000; 2000US-0636215.
PR   29-AUG-2000; 2000US-0651236.
PR   06-SEP-2000; 2000US-0657279.
PR   02-OCT-2000; 2000US-0679426.
PR   10-OCT-2000; 2000US-0685166.
XX
PA   (XUJJ/) XU J.
PA   (DILL/) DILLON D C.
PA   (MITC/) MITCHAM J L.
PA   (HARL/) HARLOCKER S L.
PA   (JIAN/) JIANG Y.
PA   (KALO/) KALOS M D.
PA   (FANG/) FANGER G R.
PA   (RETT/) RETTER M W.
PA   (STOL/) STOLK J A.
PA   (DAYC/) DAY C H.
PA   (VEDV/) VEDVICK T S.
PA   (CART/) CARTER D.
PA   (LISX/) LI S X.
PA   (WANG/) WANG A.
PA   (SKEI/) SKEIKY Y A W.
PA   (HEPL/) HEPLER W T.
PA   (HEND/) HENDERSON R A.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Kalos MD;
PI   Fanger GR,  Retter MW,  Stolk JA,  Day CH,  Vedvick TS,  Carter D;
PI   Li SX,  Wang A,  Skeiky YAW,  Hepler WT,  Henderson RA;
XX
DR   WPI; 2002-255649/30.
XX
PT   New prostate-specific polynucleotides for diagnosing and treating
PT   diseases, in particular prostate cancer, and as markers for the
PT   progression of cancer  -
XX
PS   Claim 2; SEQ ID NO 706; 87pp; English.
XX
CC   The present invention provides prostate-specific coding sequences and
CC   their encoded proteins. These can be used in the diagnosis and treatment
CC   of cancers, particularly prostate cancer. The present sequence is a
CC   protein described in the invention.
XX
SQ   Sequence   123 AA;

Alignment Scores:
Pred. No.:              6.07e-24     Length:         123
Score:                  426.00       Matches:        87
Percent Similarity:     97.83%       Conservative:   3
Best Local Similarity:  94.57%       Mismatches:     2
Query Match:            6.64%        Indels:         0
DB:                     23           Gaps:           0

US-09-759-143-110 (1-3410) x ABB95333 (1-123)

Qy       1250 ATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATG 1309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMet 20

Qy       1310 GACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTC 1369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaAlaPhe 40

Qy       1370 CCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCC 1429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAlaSerAla 60

Qy       1430 GCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTC 1489
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSerLeu 80

Qy       1490 TACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATAC 1525
              ||||||||||||||||||||   :::    ::::::
Db         81 TyrHisArgGluLysGlnValLeuIleGlyGlnTrp 92

RESULT 43
AAR80327
ID   AAR80327 standard; Protein; 762 AA.
XX
AC   AAR80327;
XX
DT   18-APR-1996  (first entry)
XX
DE   Protein polymeric adhesion substrate 1-B.
XX
KW   Pendent group; repeating unit; enzyme recognition site; sealant; fibrin;
KW   enzymatic cross-linking; biocompatible material; structural integrity;
KW   medical adhesive; wound closure; tissue repair; transglutaminase.
XX
OS   Synthetic.
XX
PN   WO9523611-A1.
XX
PD   08-SEP-1995.
XX
PF   03-MAR-1995;   95WO-US02728.
XX
PR   03-MAR-1994;   94US-0205518.
XX
PA   (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.
XX
PI   Cappello J;
XX
DR   WPI; 1995-320413/41.
XX
PT   Protein polymers comprising repeating units and sequences  - capable
PT   of enzyme-catalysed covalent bond formation useful as a
PT   biocompatible material for wound closure and tissue repair
XX
PS   Example 4; Page 45; 138pp; English.
XX
CC   The amino acid sequence of the protein polymeric adhesion substrate
CC   (PPAS) 1-B. The protein comprises 10 repeats of the PPAS1-B monomeric
CC   repeat (AAR80326) which consists of the CLP 3.7 gene encoded sequence
CC   (AAR80321) linked to the human fibrin cross-linking substrate peptide
CC   93.2 sequence (AAR80316).  The polymers can be used in biological
CC   systems where in situ formation of a biocompatible material with
CC   structural integrity is required e.g. as medical adhesives and sealants
CC   or for wound closure or tissue repair.
XX
SQ   Sequence   762 AA;

Alignment Scores:
Pred. No.:              2.42e-21     Length:         762
Score:                  395.00       Matches:        257
Percent Similarity:     34.50%       Conservative:   38
Best Local Similarity:  30.06%       Mismatches:     298
Query Match:            6.35%        Indels:         263
DB:                     16           Gaps:           52

US-09-759-143-110 (1-3410) x AAR80327 (1-762)

Qy       2347 GGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCT--- 2291
              |||  ||||||                  ||||||       |||        |||
Db         76 GlySerProGly---------------AlaProGlyThrProGlyGluGlyGlnGlnHis 90

Qy       2290 CAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCA---------GAG 2240
```

```
                 |||||||||                         :::|||:::|||    |||
Db          91 HisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaProGlyThrPro 110

Qy        2239 TCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACC 2180
               |||                    ||||||  |||  |||             |||
Db         111 GlyProGlnGlyLeuProGlySerPro---GlyAlaPro---------GlyThrProGly 126

Qy        2179 CCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCCC 2120
               |||   :::||||||   |||||||||                    |||||||||
Db         127 Pro---GlnGlyLeu---ProGlySerPro-----------GlyAlaProGlyThrPro 140

Qy        2119 CAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGT------GGCAGCAGAGAG 2066
                  |||      :::        :::      |||  |||      |||     :::
Db         141 GlyProGlnGlyLeuPro---GlySerProGlyAlaProGlyThrProGlyGluGlyGln 159

Qy        2065 CCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAAC 2006
                  |||:::   |||                                 ||||||
Db         160 GlnHisHisLeuGly-----------------------------------GlyAlaArg 167

Qy        2005 AGGAGCGGGGAGCTGGGACCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTT 1946
               ::::::|||::::::|||         |||  |||  ||||||
Db         168 GlnAlaGlyAspValGlySerPro--GlyAlaProGlyThrPro--------------- 181

Qy        1945 CTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAAT 1886
                                             |||  ||||||
Db         182 -------------------------GlyProGlnGlyLeuPro--------------- 187

Qy        1885 GGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTG 1826
                           ||||||     ||||||    |||
Db         188 -----GlySerProGlyAlaProGlyThrProGlyProGln------------------ 199

Qy        1825 GACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCACTATCCAG 1766
                                           |||  ||||||:::   |||         ||||
Db         200 -------------------------GlyLeuProGlySerProGlyAla------ProG 209

Qy        1765 GATGGCGAGGTCCAGGCAGATGCCCCGGC-----------CCGGAACCACCCTGGCCTC 1718
               ||    ||||||     ||||||            |||     ||||||
Db         209 lyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro-GlyThrProGlyGlu 228

Qy        1717 GGTGGGCTCACCCACCACCACACGTACGGAGACATCACAGGCA---------GAGGCCCC 1667
               |||          ||||||     |||       ||||||         :::||
Db         229 Gly------GlnGlnHisHisLeuGlyGly-AlaArgGlnAlaGlyAspValGlySerPr 246

Qy        1666 GCAGAGCGCGGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGG 1607
               |   :::    |||       |||    :::    ||||||    :::|||
Db         246 oGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro------------- 261

Qy        1606 GAAGGGAGCTCCAGGCTTAGGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCT 1547
                   ||||||||||||      ||||||    :::    :::   |||    |||
Db         262 ----GlyAlaProGlyThr---ProGlyProGlnGlyLeu-ProGlySerPro------- 276

Qy        1546 AGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGTAGAG 1487
                                              |||   |||   :::||||||   :::|
Db         277 ------------------------------GlyAlaProGlyThrProGlyProGlnG 286

Qy        1486 GGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGC 1427
               ||  |||            ||||||   |||       |||    ||||||||||||
Db         286 lyLeuPro-----------GlySerProGlyAla-ProGlyThrProGlyGluGly--- 300

Qy        1426 TGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACCGGCAG---------- 1377
                            ||||||              ||||||      |||
Db         301 ---GlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAla 319

Qy        1376 ---------------CCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGC 1332
                              |||||||||   ||||||    |||
Db         320 ProGlyThrProGlyProGlnGly---LeuProGlySerPro----------------- 332

Qy        1331 CGAATCGCTGCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCA 1272
                         ||||||    |||   |||:::        ||||||
Db         333 ------GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro 350

Qy        1271 GGAACA-------------------------------GCCCCAGGCTGCCCATCCGAA 1245
               ||||||                               |||||||||   |||   |||
Db         351 GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGlu 370

Qy        1244 CGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCA------CGC 1191
                      |||   :::   |||   |||                :::    :::
Db         371 GlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAla 390

Qy        1190 CCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCA 1131
               |||||||||   |||   |||:::                             :::|||
Db         391 ProGlyThr---ProGlyProGlnGly--------------------------LeuPro 400

Qy        1130 TCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGTGCGGGGCATGCGGCAGCACA 1071
                  |||      |||                      |||
Db         401 GlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 420

Qy        1070 GCTGGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCC---GGAAAGCCAAGCGGCCCGGC 1014
                  |||      |||      |||   |||||||||   |||   |||      ||||||
Db         421 ProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly 440

Qy        1013 ------------------------------------------ATGGACAGCAGTGGG 999
                                                         :::    |||   |||
Db         441 GluGlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGly 460

Qy         998 GCGACA---AGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCT 942
               |||            ||||||         |||                  |||   |||
Db         461 AlaProGlyThrProGlyProGlnGlyLeu---------ProGlySerProGlyAlaPro 477

Qy         941 CCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAA 882
                  |||          |||          |||      |||      |||
Db         478 GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro 497

Qy         881 AGAGGCACTCCT--------------CCTGGGTGCCCAGGTAGGGGGCCAGGGCACTGG 837
               :::|||   |||                  ||||||   ||||||   |||      |||
Db         498 GlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGluGlyGlnGlnHisHis 517

Qy         836 TGTCCCAGTCAATGGCAGGCAGGA------GGTAGCCCAGGCAGCCCCCAAGACTGATCA 783
                           |||||||||         ||||||||||||:::|||
Db         518 LeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaPro----------- 533

Qy         782 TGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGT---CCCGGAAGAGGT 729
                                ||||||      |||   ||||||    ||||||
Db         534 ---------------GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro 548

Qy         728 CAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCA 669
                        ||||||            :::||||||      :::|||      |||
Db         549 GlyThrProGlyProGlnGlyLeuProGlySerProGly------AlaProGlyThrPro 566

Qy         668 CGCCCAGGATGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTA 609
                  |||             :::|||||||||      |||||||||
Db         567 GlyProGlnGlyLeuProGlySerProGlyAla------ProGlyThrProGlyGluGly 584

Qy         608 GCCAGCCGGCCCTTGGGATGAGAAAGA--------------GGCTCAGCAGGATGCCCA 564
                           ||||||   |||                  ||||||   |||   |||
Db         585 GlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaPro 604

Qy         563 AGGACAGTG---CCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCAC 507
               |||         |||:::         |||
Db         605 GlyThrProGlyProGlnGlyLeuProGly------------------------------ 614

Qy         506 TGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCA 447
                     |||   ||||||         |||||||||
Db         615 ------SerProGlyAlaProGlyThrProGlyPro----------------------- 624

Qy         446 TGGTCATGAACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGC 387
                           ||||||      |||   |||
Db         625 ------------GlnGlyLeuProGlySerProGlyAla-------------------- 633

Qy         386 CTGCGGCCAAACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTT 327
                     |||   ||||||      |||   |||   |||   |||         |||
Db         634 ------ProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr 651

Qy         326 TCCGGT---------------------------------GCCGCAGCAGGCGGCTCA 303
                  |||                                 |||         ||||||
Db         652 ProGlyGluGlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySer 671
```

```
Qy        302 CCCACAGCCTCTGGACCATAGTGGGCCAGGCGGG---TAGGGCTCAGGGGGCCGTTCAGG 246
              |||  |||  |||       ||||||  |||     ||||||  |||        ||
Db        672 ProGlyAlaProGlyThr-ProGlyProGlnGlyLeuProGlySerProGlyAlaProGl 691

Qy        245 CACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTG 186
              ||||||          |||     :::|||  |||      :::|||
Db        691 yThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly----- 709

Qy        185 CCGCCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCC 126
                        |||:::      ||||||:::        |||        |||
Db        710 ---------ProGlnGlyLeuProGlySerProGlyAla---ProGlyThrProGlyGl 725

Qy        125 AGCTTCTCAGCCCATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTGGGACACGTCTC 66
                   |||    |||             ||||||  |||  |||  |||
Db        725 uGlyGlnGlnHisHisLeuGly------------GlyAlaArgGlnAlaGlyAspValGl 741

Qy         65 ATCA---------CTCAGATCCTGGCCGA 46
                |||         ::: |||||||||||
Db        741 ySerProGlyAlaMet-AspProGlyArg 750
```

RESULT 44
AAW49715
ID AAW49715 standard; Protein; 762 AA.
XX
AC AAW49715;
XX
DT 12-OCT-1998 (first entry)
XX
DE Protein polymer adhesive substrate PPAS1-B.
XX
KW Protein polymer adhesive substrate; PPAS1-B; sealant;
KW wound healing; CLP 3.7; collagen; human.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN US5773577-A.
XX
PD 30-JUN-1998.
XX
PF 03-MAR-1994; 94US-0205518.
XX
PR 02-MAR-1995; 95US-0397633.
PR 03-MAR-1994; 94US-0205518.
XX
PA (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.
XX
PI Cappello J;
XX
DR WPI; 1998-387091/33.
XX
PT New recombinant protein polymers - containing naturally occurring
PT repetitive units for crosslinking by enzymes, useful as medical
PT adhesives and sealants, depots and matrices
XX
PS Example 4; Column 27; 70pp; English.
XX
CC Protein polymer adhesive substrate PPAS1-B comprises 10 repeats
CC of a monomer (see AAW49714) consisting of human fibrin gamma chain
CC peptide 93.2 (see AAW49708) inserted into a structural backbone of
CC collagen repeat motifs (see AAW49711). The 68 kDa PPAS1-B polymer
CC has been expressed in Escherichia coli HB101 transformants using
CC plasmid pPT0324. Claimed recombinant protein polymers are capable
CC of covalent crosslinking by enzymatic reaction to form products
CC which set quickly and have good adhesive properties and high
CC strength. CLP (collagen like protein) polymers are extremely soluble
CC in water, allowing protein solutions of over 10 wt.% to be formed
CC while maintaining good flow properties. CLP polymers have good
CC adhesion to hydrophilic surfaces and therefore may adhere well to
CC tissue. The fibrin peptide 93.2 inserts of PPAS1-B provide
CC glutamine donor sites for Factor XIIIa mediated cross-linking.
XX
SQ Sequence 762 AA;

```
Alignment Scores:
Pred. No.:                2.42e-21     Length:          762
Score:                    395.00       Matches:         257
Percent Similarity:       34.50%       Conservative:    38
Best Local Similarity:    30.06%       Mismatches:      298
Query Match:              6.35%        Indels:          263
DB:                       19           Gaps:            52
```

US-09-759-143-110 (1-3410) x AAW49715 (1-762)

```
Qy       2347 GGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCT--- 2291
                |||  ||||||                ||||||      |||       |||
Db         76 GlySerProGly---------------AlaProGlyThrProGlyGluGlyGlnGlnHis 90

Qy       2290 CAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCA---------GAG 2240
                  |||||||||                       :::|||:::|||    |||
Db         91 HisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaProGlyThrPro 110

Qy       2239 TCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACC 2180
                 |||                  ||||||   |||   |||         |||
Db        111 GlyProGlnGlyLeuProGlySerPro---GlyAlaPro---------GlyThrProGly 126

Qy       2179 CCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCCC 2120
              |||   :::||||||        |||||||||                 |||||||||
Db        127 Pro--GlnGlyLeu---ProGlySerPro------------GlyAlaProGlyThrPro 140

Qy       2119 CAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGT------GGCAGCAGAGAG 2066
                 |||      :::          :::     |||   |||      |||     :::
Db        141 GlyProGlnGlyLeuPro---GlySerProGlyAlaProGlyThrProGlyGluGlyGln 159

Qy       2065 CCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAAC 2006
                 |||:::   |||                                   ||||||
Db        160 GlnHisHisLeuGly-----------------------------------GlyAlaArg 167

Qy       2005 AGGAGCGGGGAGCTGGGACCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTT 1946
              ::::::|||::::::|||            |||     |||     ||||||
Db        168 GlnAlaGlyAspValGlySerPro--GlyAlaProGlyThrPro--------------- 181

Qy       1945 CTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAAT 1886
                                            |||  ||||||
Db        182 ------------------------GlyProGlnGlyLeuPro---------------- 187

Qy       1885 GGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTG 1826
                          ||||||           ||||||     |||
Db        188 -----GlySerProGlyAlaProGlyThrProGlyProGln------------------ 199

Qy       1825 GACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCACTATCCAG 1766
                                            |||   ||||||:::   |||       ||||
Db        200 ------------------------GlyLeuProGlySerProGlyAla------ProG 209

Qy       1765 GATGGCGAGGTCCAGGCAGATGCCCCGGC-----------CCGGAACCACCCTGGCCTC 1718
              ||     ||||||               ||||||           |||       ||||||
Db        209 lyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro-GlyThrProGlyGlu 228

Qy       1717 GGTGGGCTCACCCACCACCACACGTACGGAGACATCACAGGCA---------GAGGCCCC 1667
              |||         ||||||     |||        ||||||           :::||
Db        229 Gly------GlnGlnHisHisLeuGlyGly-AlaArgGlnAlaGlyAspValGlySerPr 246

Qy       1666 GCAGAGCGCGGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGG 1607
              |    :::   |||        |||    :::    ||||||    :::|||
Db        246 oGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro------------- 261

Qy       1606 GAAGGGAGCTCCAGGCTTAGGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCT 1547
                  ||||||||||||         ||||||     :::    :::    |||   |||
Db        262 ----GlyAlaProGlyThr---ProGlyProGlnGlyLeu-ProGlySerPro------- 276

Qy       1546 AGCACCTCCAGTGTCCCCTCGGTATTTGGCAGGAACACCTGCTTCTCCCGGTGGTAGAG 1487
                                              |||    |||    :::||||||    :::|
Db        277 ------------------------------GlyAlaProGlyThrProGlyProGlnG 286

Qy       1486 GGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGC 1427
```

```
                  ||   |||                ||||||   |||        |||   |||||||||||
Db        286 lyLeuPro-----------GlySerProGlyAla-ProGlyThrProGlyGluGly--- 300

Qy       1426 TGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACCGGCAG---------- 1377
                          ||||||                 ||||||        |||
Db        301 ---GlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAla 319

Qy       1376 ---------------CCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGC 1332
                             |||||||||  ||||||       |||
Db        320 ProGlyThrProGlyProGlnGly---LeuProGlySerPro------------------ 332

Qy       1331 CGAATCGCTGCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCA 1272
                        ||||||     |||   |||:::      ||||||
Db        333 ------GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro 350

Qy       1271 GGAACA--------------------------------GCCCCAGGCTGCCCATCCGAA 1245
              ||||||                                |||||||||   |||   |||
Db        351 GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGlu 370

Qy       1244 CGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCA------CGC 1191
                        |||   :::   |||   |||                :::   :::
Db        371 GlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAla 390

Qy       1190 CCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCA 1131
              |||||||||   |||   |||:::                                :::|||
Db        391 ProGlyThr---ProGlyProGlnGly-------------------------LeuPro 400

Qy       1130 TCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACA 1071
                   |||      |||                      |||
Db        401 GlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 420

Qy       1070 GCTGGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCC---GGAAAGCCAAGCGGGCCCGGC 1014
                 |||      |||      |||   |||||||||   |||   |||       ||||||
Db        421 ProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly 440

Qy       1013 ---------------------------------------------ATGGACAGCAGTGGG 999
                                                           :::   |||   |||
Db        441 GluGlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGly 460

Qy        998 GCGACA---AGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCT 942
              |||         ||||||     |||                      |||    |||
Db        461 AlaProGlyThrProGlyProGlnGlyLeu---------ProGlySerProGlyAlaPro 477

Qy        941 CCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAA 882
                    |||            |||         |||      |||      |||
Db        478 GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro 497

Qy        881 AGAGGCACTCCT---------------CCTGGGTGCCCAGGTAGGGGGCCAGGGCACTGG 837
              :::|||    |||                ||||||   ||||||    |||      |||
Db        498 GlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGluGlyGlnGlnHisHis 517

Qy        836 TGTCCCAGTCAATGGCAGGCAGGA------GGTAGCCCAGGCAGCCCCCAAGACTGATCA 783
                          |||||||||      |||||||||||:::|||
Db        518 LeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaPro------------ 533

Qy        782 TGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGT---CCCGGAAGAGGT 729
                                      ||||||   |||   ||||||   ||||||
Db        534 ---------------GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro 548

Qy        728 CAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCA 669
                    ||||||                 :::||||||      :::|||      |||
Db        549 GlyThrProGlyProGlnGlyLeuProGlySerProGly------AlaProGlyThrPro 566

Qy        668 CGCCCAGGATGAGCAGTGCCAGCTCCAGGGGCCTGGATCCGGGCACAGCAGCCCTGCTA 609
                 |||                :::|||||||||      |||||||||
Db        567 GlyProGlnGlyLeuProGlySerProGlyAla------ProGlyThrProGlyGluGly 584

Qy        608 GCCAGCCGGCCCTTGGGATGAGAAAGA---------------GGCTCAGCAGGATGCCCA 564
                       ||||||   |||                  ||||||   |||   |||
Db        585 GlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaPro 604

Qy        563 AGGACAGTG---CCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCAC 507
                 |||      |||:::          |||

Db        605 GlyThrProGlyProGlnGlyLeuProGly---------------------------- 614

Qy        506 TGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCA 447
                    |||   ||||||          |||||||||
Db        615 ------SerProGlyAlaProGlyThrProGlyPro---------------------- 624

Qy        446 TGGTCATGAACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGC 387
                            ||||||   |||   |||
Db        625 ------------GlnGlyLeuProGlySerProGlyAla--------------------- 633

Qy        386 CTGCGGCCAAACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTT 327
                    |||   ||||||      |||   |||   |||   |||         |||
Db        634 ------ProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr 651

Qy        326 TCCGGT------------------------------------GCCGCAGCAGGCGGCTCA 303
                 |||                                |||         ||||||
Db        652 ProGlyGluGlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySer 671

Qy        302 CCCACAGCCTCTGGACCATAGTGGCCAGGCGGG---TAGGGCTCAGGGGGCCGTTCAGG 246
              |||   |||   |||      ||||||   |||      ||||||   |||      ||
Db        672 ProGlyAlaProGlyThr-ProGlyProGlnGlyLeuProGlySerProGlyAlaProGl 691

Qy        245 CACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTG 186
              |||||||        |||      :::|||   |||        :::|||
Db        691 yThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly----- 709

Qy        185 CCGCCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCC 126
                          |||:::     ||||||:::                |||       |||
Db        710 ----------ProGlnGlyLeuProGlySerProGlyAla---ProGlyThrProGlyGl 725

Qy        125 AGCTTCTCAGCCCATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTC 66
                      |||      |||                ||||||   |||   |||   |||
Db        725 uGlyGlnGlnHisHisLeuGly-----------GlyAlaArgGlnAlaGlyAspValGl 741

Qy         65 ATCA---------CTCAGATCCTGGCCGA 46
                |||           :::  ||||||||||||
Db        741 ySerProGlyAlaMet-AspProGlyArg 750
```

```
RESULT 45
AAW57668
ID   AAW57668 standard; peptide; 762 AA.
XX
AC   AAW57668;
XX
DT   27-AUG-1998  (first entry)
XX
DE   Collagen-like polymer.
XX
KW   Collagen-like polymer; synthetic polymer; fibre coating;
KW   prosthetic device; catalytic substance.
XX
OS   Synthetic.
XX
PN   US5773249-A.
XX
PD   30-JUN-1998.
XX
PF   02-MAY-1996;   96US-0642255.
XX
PR   02-MAY-1996;   96US-0642255.
PR   04-NOV-1986;   86US-0927258.
PR   29-OCT-1987;   87US-0114618.
PR   09-NOV-1988;   88US-0269429.
PR   06-NOV-1990;   90US-0609716.
PR   12-NOV-1991;   91US-0791960.
PR   05-NOV-1992;   92US-0972032.
PR   22-DEC-1995;   95US-0577046.
XX
PA   (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.
XX
PI   Cappello J,  Ferrari FA;
XX
DR   WPI; 1998-387004/33.
```

```
XX
PT   Recombinant collagen-like polymers - useful for making gels, films,
PT   fibres, etc.
XX
PS   Example 7; Column 51; 93pp; English.
XX
CC   This sequence represents a unnatural collagen-like polymer
CC   of the invention. The products may be used as films, fibres, moulded
CC   objects and admixed with other natural or synthetic polymers or coatings
CC   on fibres, films, labware or other surfaces, e.g. prosthetic devices. The
CC   polymers may be used for binding a wide variety of specific binding
CC   materials, as catalytic substances (where the amino acid sequence may
CC   specifically chelate a wide variety of elements), as purification media,
CC   composites, laminates or adhesives. They may also be combined with
CC   inorganic or organic materials such as carbon fibres, nylon fibres,
CC   nitrocellulose, etc., as flask coatings or in synthetic matrices for the
CC   growth and study of cells, as affinity columns or as supports for
CC   biological materials. The polymers have collagen-like properties, but may
CC   be easily expressed in micro-organisms in high efficiency. The new
CC   sequences can be tailored to give the desired properties.
XX
SQ   Sequence   762 AA;

Alignment Scores:
Pred. No.:              2.42e-21        Length:         762
Score:                  395.00          Matches:        257
Percent Similarity:     34.50%          Conservative:   38
Best Local Similarity:  30.06%          Mismatches:     298
Query Match:            6.35%           Indels:         263
DB:                     19              Gaps:           52

US-09-759-143-110 (1-3410) x AAW57668 (1-762)

Qy     2347 GGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCT--- 2291
              |||  ||||||                   ||||||     |||    |||
Db       76 GlySerProGly---------------AlaProGlyThrProGlyGluGlyGlnGlnHis 90

Qy     2290 CAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCA---------GAG 2240
                |||||||||                    :::|||:::|||  |||
Db       91 HisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaProGlyThrPro 110

Qy     2239 TCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACC 2180
                |||                 ||||||   |||   |||              |||
Db      111 GlyProGlnGlyLeuProGlySerPro---GlyAlaPro---------GlyThrProGly 126

Qy     2179 CCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCCC 2120
             |||   :::||||||      |||||||||                   |||||||||
Db      127 Pro---GlnGlyLeu---ProGlySerPro-----------GlyAlaProGlyThrPro 140

Qy     2119 CAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGT------GGCAGCAGAGAG 2066
             |||      :::            :::      |||    |||    |||       :::
Db      141 GlyProGlnGlyLeuPro---GlySerProGlyAlaProGlyThrProGlyGluGlyGln 159

Qy     2065 CCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAAC 2006
               |||:::   |||                                   ||||||
Db      160 GlnHisHisLeuGly-----------------------------------GlyAlaArg 167

Qy     2005 AGGAGCGGGGAGCTGGGACCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTT 1946
             ::::::|||::::::|||           |||    |||    ||||||
Db      168 GlnAlaGlyAspValGlySerPro--GlyAlaProGlyThrPro---------------- 181

Qy     1945 CTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAAT 1886
                                              |||   ||||||
Db      182 -------------------------GlyProGlnGlyLeuPro---------------- 187

Qy     1885 GGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTG 1826
                               ||||||        ||||||    |||
Db      188 -----GlySerProGlyAlaProGlyThrProGlyProGln------------------- 199

Qy     1825 GACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCACTATCCAG 1766
                                          |||   ||||||:::  |||         ||||
Db      200 -------------------------GlyLeuProGlySerProGlyAla------ProG 209

Qy     1765 GATGGCGAGGTCCAGGCAGATGCCCCGGC-----------CCGGAACCACCCTGGCCTC 1718
             ||      ||||||          ||||||            |||      ||||||
Db      209 lyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro-GlyThrProGlyGlu 228

Qy     1717 GGTGGGCTCACCCACCACCACACGTACGGAGACATCACAGGCA---------GAGGCCCC 1667
             |||            ||||||   |||          ||||||          :::||
Db      229 Gly------GlnGlnHisHisLeuGlyGly-AlaArgGlnAlaGlyAspValGlySerPr 246

Qy     1666 GCAGAGCGCGGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGG 1607
             |   :::    |||          |||     :::     ||||||    :::|||
Db      246 oGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro-------------- 261

Qy     1606 GAAGGGAGCTCCAGGCTTAGGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCT 1547
                 |||||||||||       ||||||    :::    :::     |||    |||
Db      262 ----GlyAlaProGlyThr---ProGlyProGlnGlyLeu-ProGlySerPro------- 276

Qy     1546 AGCACCTCCAGTGTCCCCTCGGTATTTGGCAGGAACACCTGCTTCTCCCGGTGGTAGAG 1487
                                             |||   |||    :::||||||    :::|
Db      277 --------------------------------GlyAlaProGlyThrProGlyProGlnG 286

Qy     1486 GGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGC 1427
             ||   |||            ||||||   |||        |||    ||||||||||||
Db      286 lyLeuPro-----------GlySerProGlyAla-ProGlyThrProGlyGluGly--- 300

Qy     1426 TGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACCGGCAG---------- 1377
                    ||||||                  ||||||        |||
Db      301 ---GlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAla 319

Qy     1376 ---------------CCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGC 1332
                            |||||||||   ||||||        |||
Db      320 ProGlyThrProGlyProGlnGly---LeuProGlySerPro----------------- 332

Qy     1331 CGAATCGCTGCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCA 1272
                        ||||||     |||    |||:::     ||||||
Db      333 ------GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro 350

Qy     1271 GGAACA-------------------------------GCCCCAGGCTGCCCATCCGAA 1245
             ||||||                                ||||||||||   |||    |||
Db      351 GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGlu 370

Qy     1244 CGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCA------CGC 1191
                      |||    :::    |||    |||                 :::    :::
Db      371 GlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAla 390

Qy     1190 CCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCA 1131
             |||||||||   |||    |||:::                                  :::|||
Db      391 ProGlyThr---ProGlyProGlnGly-------------------------LeuPro 400

Qy     1130 TCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGTGCGGGGCATGCGGCAGCACA 1071
                |||         |||                       |||
Db      401 GlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 420

Qy     1070 GCTGGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCC---GGAAAGCCAAGCGGGCCCGGC 1014
                |||       |||       |||     |||||||||     |||    |||       ||||||
Db      421 ProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly 440

Qy     1013 --------------------------------------------ATGGACAGCAGTGGG 999
                                                           :::    |||    |||
Db      441 GluGlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGly 460

Qy      998 GCGACA---AGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCT 942
             |||                ||||||     |||                     |||     |||
Db      461 AlaProGlyThrProGlyProGlnGlyLeu---------ProGlySerProGlyAlaPro 477

Qy      941 CCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAA 882
                   |||              |||              |||     |||     |||
Db      478 GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro 497

Qy      881 AGAGGCACTCCT---------------CCTGGGTGCCCAGGTAGGGGGCCAGGGCACTGG 837
             :::|||    |||                  ||||||  ||||||    |||      |||
Db      498 GlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGluGlyGlnGlnHisHis 517

Qy      836 TGTCCCAGTCAATGGCAGGCAGGA------GGTAGCCCAGGCAGCCCCCAAGACTGATCA 783
```

```
                 |||||||||       ||||||||||||:::|||
Db     518 LeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaPro----------- 533

Qy     782 TGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGT---CCCGGAAGAGGT 729
                                  ||||||     |||  ||||||   ||||||
Db     534 ---------------GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro 548

Qy     728 CAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCA 669
                  ||||||                :::||||||      :::|||       |||
Db     549 GlyThrProGlyProGlnGlyLeuProGlySerProGly------AlaProGlyThrPro 566

Qy     668 CGCCCAGGATGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTA 609
              |||                 :::|||||||||      |||||||||
Db     567 GlyProGlnGlyLeuProGlySerProGlyAla------ProGlyThrProGlyGluGly 584

Qy     608 GCCAGCCGGCCCTTGGGATGAGAAAGA---------------GGCTCAGCAGGATGCCCA 564
                    ||||||     |||                 ||||||   |||   |||
Db     585 GlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaPro 604

Qy     563 AGGACAGTG---CCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCAC 507
              |||      |||:::          |||
Db     605 GlyThrProGlyProGlnGlyLeuProGly---------------------------- 614

Qy     506 TGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCA 447
                 |||   ||||||        ||||||||
Db     615 ------SerProGlyAlaProGlyThrProGlyPro---------------------- 624

Qy     446 TGGTCATGAACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGC 387
                             ||||||      |||   |||
Db     625 -----------GlnGlyLeuProGlySerProGlyAla-------------------- 633

Qy     386 CTGCGGCCAAACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTT 327
                 |||  ||||||       |||   |||   |||   |||       |||
Db     634 ------ProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr 651

Qy     326 TCCGGT----------------------------------GCCGCAGCAGGCGGCTCA 303
              |||                                   |||        ||||||
Db     652 ProGlyGluGlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySer 671

Qy     302 CCCACAGCCTCTGGACCATAGTGGGCCAGGCGGG---TAGGGCTCAGGGGGCCGTTCAGG 246
            |||   |||   |||        ||||||   |||      ||||||   |||      ||
Db     672 ProGlyAlaProGlyThr-ProGlyProGlnGlyLeuProGlySerProGlyAlaProGl 691

Qy     245 CACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTG 186
            |||||||                |||       :::|||   |||      :::|||
Db     691 yThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly-----709

Qy     185 CCGCCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCC 126
                          |||:::      ||||||:::         |||      |||
Db     710 ----------ProGlnGlyLeuProGlySerProGlyAla---ProGlyThrProGlyGl 725

Qy     125 AGCTTCTCAGCCCATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTC 66
                 |||      |||                ||||||   |||   |||   |||
Db     725 uGlyGlnGlnHisHisLeuGly-----------GlyAlaArgGlnAlaGlyAspValGl 741

Qy      65 ATCA---------CTCAGATCCTGGCCGA 46
            |||         ::: ||||||||||||
Db     741 ySerProGlyAlaMet-AspProGlyArg 750

Search completed: June  9, 2003, 22:24:42
Job time : 166.5 secs
```

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         June  9, 2003, 13:07:53 ; Search time 709 Seconds
                                          (without alignments)
                                          10831.191 Million cell updates/sec

Title:          US-09-759-143-110
Perfect score:  3410
Sequence:       1 gggaaccagcctgcacgcgc..........aaaaaaataaaaaaaaaaaa 3410

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters:      4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       N_Geneseq_101002:*
                 1:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
                 2:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
                 3:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
                 4:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
                 5:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
                 6:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
                 7:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
                 8:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
                 9:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
                 10:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
                 11:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
                 12:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
                 13:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
                 14:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
                 15:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
                 16:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
                 17:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
                 18:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
                 19:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
                 20:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
                 21:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
                 22:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
                 23:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
                 24:  /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                 SUMMARIES
                   %
Result           Query
  No.   Score    Match Length DB ID                  Description
-----------------------------------------------------------------------
    1  3409.6   100.0   3410  19 AAV61201            Full length cDNA s
    2  3409.6   100.0   3410  19 AAV58586            Prostate tumour sp
    3  3409.6   100.0   3410  21 AAC79473            cDNA sequence of h
    4  3409.6   100.0   3410  22 AAS63557            Human prostate cDN
    5  3409.6   100.0   3410  22 AAS10108            Human prostate tum
    6  3409.6   100.0   3410  22 AAH93465            Human prostate-spe
    7  3409.6   100.0   3410  22 AAH84779            Human prostate-spe
    8  3409.6   100.0   3410  22 AAH02530            Prostate tumour an
    9  3409.6   100.0   3410  24 ABL94929            Human L1-12 cDNA s
   10  3409.6   100.0   3410  24 ABK29017            Human breast tumou
   11    3408    99.9   3410  21 AAA06349            Human immunogenic
   12  3292.4    96.6   3320  24 AAS14962            DNA encoding human
   13  2585.4    75.8   4034  22 AAS64040            Human prosate cDNA
   14  2585.4    75.8   4034  22 AAH93868            P553S cDNA splice
   15  2585.4    75.8   4034  24 ABL95411            Human P553S splice
   16  2486.4    72.9   2582  24 ABK92217            Prostate cancer-as
   17  2196.4    64.4   2904  22 AAS64039            Human prosate cDNA
   18  2196.4    64.4   2904  22 AAH93867            P553S cDNA splice
   19  2196.4    64.4   2904  24 ABL95410            Human P553S splice
   20  2142.8    62.8   4894  22 AAS64038            Human prosate cDNA
   21  2142.8    62.8   4894  22 AAH93866            P553S cDNA splice
   22  2142.8    62.8   4894  24 ABL95409            Human P553S splice
   23  2136.8    62.7   3878  22 AAD05230            Human secreted pro
   24  2136.4    62.7   2152  20 AAV71181            Consensus sequence
   25  2114.8    62.0   2143  20 AAV71180            Clone 1711346IH, t
   26  2065.4    60.6   2462  21 AAZ45677            cDNA sequence of a
   27    1915    56.2   2133  21 AAC64928            Human prostate-rel
   28    1915    56.2   2133  22 AAS07155            Prostate gene PS10
   29  1899.8    55.7   2124  21 AAC64927            Human prostate-rel
   30  1899.8    55.7   2124  22 AAS07601            Prostate gene PS10
   31  1854.8    54.4   3663  24 ABN81320            Human mast cell re
   32  1815.8    53.2   6976  22 AAS64041            Human prosate cDNA
   33  1815.8    53.2   6976  22 AAH93869            P553S cDNA splice
   34  1815.8    53.2   6976  24 ABL95412            Human P553S splice
   35    1662    48.7   1662  24 ABN81324            Human mast cell re
   36    1503    44.1   2133  24 ABA91283            Thioredoxin-ubiqui
   37   961.2    28.2   1593  24 ABA91284            Thioredoxin-ubiqui
c  38   925.2    27.1   1019  22 AAD05264            Human secreted pro
   39   794.6    23.3   1203  22 AAS64153            Human /M. tubercul
   40   794.6    23.3   1203  22 AAH93917            Ra12-P501S-E2 cons
   41   794.6    23.3   1203  24 ABL95524            Ra12-P501S-E2 cons
   42   673.4    19.7    789  19 AAV61144            3' cDNA sequence o
   43   673.4    19.7    789  19 AAV58487            3' fragment of pro
   44   673.4    19.7    789  21 AAA06250            Human immunogenic
   45   673.4    19.7    789  22 AAS63458            Human prostate cDN

                               ALIGNMENTS

RESULT 1
AAV61201
ID   AAV61201 standard; cDNA; 3410 BP.
XX
AC   AAV61201;
XX
DT   06-JAN-1999  (first entry)
XX
DE   Full length cDNA sequence of prostate tumour clone L1-12.
XX
KW   Prostate; cancer; tumour; vaccine; immunogen; clone; ss.
XX
OS   Homo sapiens.
XX
PN   WO9837093-A2.
XX
PD   27-AUG-1998.
XX
PF   25-FEB-1998;   98WO-US03492.
XX
PR   09-FEB-1998;   98US-0020956.
PR   25-FEB-1997;   97US-0806099.
PR   01-AUG-1997;   97US-0904804.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Dillon DC,  Xu J;
XX
DR   WPI; 1998-609886/51.
DR   P-PSDB; AAW71869.
XX
PT   Polypeptides comprising immunogenic portions of prostate proteins -
```

```
PT   used in a vaccine for the treatment of prostate cancer
XX
PS   Claim 3; Page 79-80; 130pp; English.
XX
CC   The present sequence is a new DNA which encodes an immunogenic portion
CC   of a prostate tumour protein. The encoded immunogen, or the DNA itself,
CC   can be used as a vaccine for the treatment of prostate cancer. The DNA
CC   was identified by analysis of a subtracted cDNA library obtained by
CC   subtracting a prostate tumour cDNA expression library with a normal
CC   tissue cDNA library.
XX
SQ   Sequence 3410 BP; 667 A; 1014 C; 945 G; 783 T; 1 other;

  Query Match             100.0%;  Score 3409.6;  DB 19;  Length 3410;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy       1741 GGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy         1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy         2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy         2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy         2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy         2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy         2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy         2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy         2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy         2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy         2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy         2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy         2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700

Qy         2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy         2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy         2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy         2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy         2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy         3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy         3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy         3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy         3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy         3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy         3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy         3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAA 3410
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAA 3410

RESULT 2
AAV58586
ID   AAV58586 standard; cDNA; 3410 BP.
XX
AC   AAV58586;
XX
DT   08-DEC-1998  (first entry)
XX
DE   Prostate tumour specific gene clone L1-12.
XX
KW   Prostate tumour specific gene; human; prostate cancer; detection;
KW   therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             284..1945
FT                   /*tag=  a
XX
PN   WO9837418-A2.
XX
PD   27-AUG-1998.
XX
PF   25-FEB-1998;   98WO-US03690.
XX
PR   09-FEB-1998;   98US-0904809.
PR   25-FEB-1997;   97US-0806596.
PR   01-AUG-1997;   97US-0904809.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Dillon DC,  Xu J;
XX
DR   WPI; 1998-480805/41.
DR   P-PSDB; AAW69385.
XX
PT   Novel human prostate specific tumour protein and fragments - useful
PT   for detecting and treating prostate cancers
XX
PS   Claim 1; Page 84-85; 141pp; English.
XX
CC   This sequence represents a human prostate tumour specific gene, and can
CC   be used in the method of the invention. The method is for detecting
CC   prostate cancer comprises contacting a biological sample with an agent
CC   able to bind an immunogenic portion of a prostate protein (such as
CC   encoded by this sequence). An antibody which binds to an immunogenic
```

```
CC   portion of the prostate protein, and the method can be used to detect,
CC   monitor progression of, or treat prostate cancers. The antibody may
CC   also be conjugated to a therapeutic agent for use in therapy of prostate
CC   cancers.
XX
SQ   Sequence 3410 BP; 667 A; 1014 C; 945 G; 783 T; 1 other;

  Query Match             100.0%;  Score 3409.6;  DB 19;  Length 3410;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
```

```
Qy       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy       2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy       2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy       2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy       2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700

Qy       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
```

```
RESULT 3
AAC79473
ID   AAC79473 standard; cDNA; 3410 BP.
XX
AC   AAC79473;
XX
DT   07-FEB-2001  (first entry)
XX
DE   cDNA sequence of human breast tumour clone P501S.
XX
KW   Human; breast tumour antigen; cytostatic; immunotherapy;
KW   breast cancer; vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200061756-A2.
XX
PD   19-OCT-2000.
XX
PF   10-APR-2000; 2000WO-US09688.
XX
PR   09-APR-1999;   99US-0288950.
PR   02-JUL-1999;   99US-0346327.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Reed SG,  Xu J,  Dillon DC;
XX
DR   WPI; 2000-638568/61.
DR   P-PSDB; AAB28527.
XX
PT   A novel isolated polypeptide comprising an immunogenic portion of a
PT   breast cancer protein useful in the detection and treatment of breast
PT   cancer -
XX
PS   Claim 26; Page 91-92; 95pp; English.
XX
CC   The present sequence was isolated from a breast tumour cDNA library. It
CC   is provided in a specification relating to compounds for immunotherapy
CC   and diagnosis of breast cancer. Breast tumour antigens and the
CC   polynucleotides that encode them may be used in the production of a
CC   pharmaceutical composition to be used in the treatment of breast cancer.
CC   Proliferated T cells and incubated antigen presenting cells are also
CC   required. The polypeptides and polynucleotides may also be used to
CC   produce a vaccine.
XX
SQ   Sequence 3410 BP; 667 A; 1014 C; 945 G; 783 T; 1 other;

  Query Match            100.0%;  Score 3409.6;  DB 21;  Length 3410;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy       2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGCTCC 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy     2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy     2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy     2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy     2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy     2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy     2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy     2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy     2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700

Qy     2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy     2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy     2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy     2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy     2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy     3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy     3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy     3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy     3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy     3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy     3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy     3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
```

```
RESULT 4
AAS63557
ID   AAS63557 standard; cDNA; 3410 BP.
XX
AC   AAS63557;
XX
DT   29-JAN-2002  (first entry)
XX
DE   Human prostate cDNA sequence #109.
XX
KW   Human; prostate cancer; ss; cytostatic; immunostimulant; tumour.
XX
OS   Homo sapiens.
XX
PN   WO200173032-A2.
XX
PD   04-OCT-2001.
XX
PF   27-MAR-2001; 2001WO-US09919.
XX
PR   27-MAR-2000; 2000US-0536857.
PR   09-MAY-2000; 2000US-0568100.
PR   12-MAY-2000; 2000US-0570737.
PR   13-JUN-2000; 2000US-0593793.
PR   27-JUN-2000; 2000US-0605783.
PR   10-AUG-2000; 2000US-0636215.
PR   29-AUG-2000; 2000US-0651236.
PR   06-SEP-2000; 2000US-0657279.
PR   02-OCT-2000; 2000US-0679426.
PR   10-OCT-2000; 2000US-0685166.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Kalos MD;
PI   Fanger GR,  Retter MW,  Stolk JA,  Day CH,  Vedvick TS,  Carter D;
PI   Li SX,  Wang A,  Skeiky YAW,  Hepler WT,  Henderson RA;
XX
DR   WPI; 2001-639232/73.
DR   P-PSDB; AAU69763.
XX
PT   New human prostate-specific polypeptides and polynucleotides useful for
PT   the diagnosis and treatment of cancer, especially prostate cancer -
XX
PS   Claim 1; Page 267-268; 579pp; English.
XX
CC   The invention relates to isolated prostate-specific
CC   polynucleotides, polypeptides, fusion proteins of the polypeptides,
CC   antibodies raised against the polypeptides (or antigenic epitopes
CC   derived from them) and antigen-presenting cells expressing the
CC   polypeptides. The antibodies are useful for detecting the presence of
CC   cancer, especially prostate cancer. The polypeptides, polynucleotides and
CC   the antigen-presenting cells are useful for stimulating and/or expanding
CC   T cells specific for a tumour protein, and for inhibiting the development
CC   of cancer especially prostate cancer. Compositions comprising the
CC   polynucleotide and/or polypeptide are useful for stimulating an immune
CC   response, and for treating cancer. The oligonucleotide is useful for
CC   detecting cancer. The present sequence is a prostate specific
CC   polynucleotide of the invention.
XX
SQ   Sequence 3410 BP; 667 A; 1014 C; 945 G; 783 T; 1 other;

  Query Match             100.0%;  Score 3409.6;  DB 22;  Length 3410;
  Best Local Similarity   100.0%;  Pred. No. 0;
```

```
Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
```

```
Qy     2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy     2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy     2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy     2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy     2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy     2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy     2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy     2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy     2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700

Qy     2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy     2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy     2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy     2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy     2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy     3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy     3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy     3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy     3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy     3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy     3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy     3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAA 3410
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAA 3410


RESULT 5
AAS10108
ID   AAS10108 standard; cDNA; 3410 BP.
XX
AC   AAS10108;
XX
DT   24-OCT-2001  (first entry)
XX
DE   Human prostate tumour cDNA L1-12.
XX
KW   Human; prostate tumour protein; prostate cancer; ss.
XX
OS   Homo sapiens.
XX
PN   US6262245-B1.
XX
PD   17-JUL-2001.
XX
PF   25-FEB-1998;   98US-0030607.
XX
PR   25-FEB-1997;   97US-0806099.
PR   01-AUG-1997;   97US-0904804.
PR   09-FEB-1998;   98US-0020956.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC;
XX
DR   WPI; 2001-440862/47.
DR   P-PSDB; AAU04961.
XX
PT   Novel polynucleotide encoding polypeptide comprising a portion of
PT   prostate tumour protein useful for inhibiting development of prostate
PT   cancer or for treating prostate cancer in a patient  -
XX
PS   Claim 3; Column 118-121; 105pp; English.
XX
CC   The sequence is a human prostate tumour cDNA which encodes a
CC   partial tumour protein. The DNA is useful for inhibiting the development
CC   of prostate cancer or for treating prostate cancer in a patient.
XX
SQ   Sequence 3410 BP; 667 A; 1014 C; 945 G; 783 T; 1 other;

  Query Match             100.0%;  Score 3409.6;  DB 22;  Length 3410;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy       61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy      121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy      181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy      241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy      301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy      361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy      421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy      481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy      541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy      601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy      661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy      721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy      781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy      841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy      901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy      961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy     1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy     1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy     1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy     1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy     1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy     1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy     1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy     1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy     1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy     1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy     1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy     1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy     1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy     1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy     1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy     1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy     1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy     2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy     2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy     2161 ACTGGAGGCCTTCCAAGGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2161 ACTGGAGGCCTTCCAAGGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy     2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy     2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy     2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
```

```
Qy       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700

Qy       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
```

RESULT 6

```
AAH93465
ID   AAH93465 standard; cDNA; 3410 BP.
XX
AC   AAH93465;
XX
DT   04-OCT-2001  (first entry)
XX
DE   Human prostate-specific full length cDNA sequence L1-12.
XX
KW   Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW   cytostatic; gene therapy; metastasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200151633-A2.
XX
PD   19-JUL-2001.
XX
PF   16-JAN-2001; 2001WO-US01574.
XX
PR   14-JAN-2000; 2000US-0483672.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Reed SG;
PI   Kalos MD,  Fanger GR,  Day CH,  Retter MW,  Stolk JA,  Skeiky YAW;
PI   Wang A,  Meagher MJ;
XX
DR   WPI; 2001-425873/45.
XX
PT   New polynucleotide encoding a prostate-specific protein, for
PT   diagnosing, monitoring and treating prostate cancer in a patient and
PT   for use in vaccines -
XX
PS   Claim 1; Page 265-266; 543pp; English.
XX
CC   The present invention describes polynucleotide sequences (I) which encode
CC   prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC   and can be used in vaccine production and gene therapy. (I), (II),
CC   antibodies to (II), fusion proteins comprising (II), and isolated
CC   T cells prepared using (I) or (II) are used treat cancer in a patient.
CC   (I) and the antibodies are also used in the detection of cancer in a
CC   patient. The cancer that is diagnosed or treated is particularly
CC   prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
CC   (I) can be used for monitoring the progression of cancer in a patient.
CC   (I) and (II) can also be used to improve diagnostic and therapeutic
CC   methods for prostate cancer. They can indicate the level of metastasis
CC   as well as the prostate volume. AAH93357 to AAH93944 and AAM01115 to
CC   AAM01318 represent polynucleotide and amino acid sequences used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 3410 BP; 667 A; 1014 C; 945 G; 783 T; 1 other;

  Query Match             100.0%;  Score 3409.6;  DB 22;  Length 3410;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
```

```
Qy    301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy    361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy    421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy    481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy    541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy    601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy    661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy    721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy    781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy    841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy    901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy    961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy   1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy   1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy   1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy   1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy   1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy   1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy   1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy   1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy   1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy   1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy   1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy   1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy   1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy   1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy   1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy   1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy   1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy   2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy   2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy   2161 ACTGGAGGCCTTCCAAGGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 ACTGGAGGCCTTCCAAGGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy   2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy   2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy   2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy   2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy   2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy         2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy         2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy         2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700

Qy         2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy         2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy         2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy         2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy         2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy         3001 CCCAACTTTCCCCTACCCCAACTTTCCCGACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3001 CCCAACTTTCCCCTACCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy         3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy         3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy         3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy         3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy         3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy         3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
```

```
RESULT 7
AAH84779
ID   AAH84779 standard; cDNA; 3410 BP.
XX
AC   AAH84779;
XX
DT   25-SEP-2001  (first entry)
XX
DE   Human prostate-specific cDNA sequence L1-12/P501S.
XX
KW   Human; prostate cancer; therapy; diagnosis; cat eye syndrome;
KW   chromosome 22q11.2; prostate-specific protein; chromosome 1;
KW   prostate specific antigen; PSA; ss.
XX
OS   Homo sapiens.
XX
PN   WO200134802-A2.
XX
PD   17-MAY-2001.
XX
PF   09-NOV-2000; 2000WO-US30904.
XX
PR   12-NOV-1999;   99US-0439313.
PR   18-NOV-1999;   99US-0443686.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Reed SG;
PI   Kalos MD,  Retter MW,  Stolk JA,  Day CH,  Skeiky YAW,  Wang A;
XX
DR   WPI; 2001-308785/32.
XX
PT   Isolated polypeptide comprising at least an immunogenic portion of a
PT   prostate-specific protein, useful in the diagnosis and therapy of
PT   prostate cancer -
XX
PS   Claim 5; Page 164-165; 325pp; English.
XX
CC   The present invention describes an isolated polypeptide (P1) comprising
CC   at least an immunogenic portion of a prostate-specific protein, or its
CC   variant. Also described are polynucleotides (N1) encoding (P1). (P1) and
CC   (N1) have cytostatic activity and can be used in vaccine production.
CC   The polypeptides, nucleic acids and antibodies from the present
CC   invention are useful in the diagnosis and therapy of prostate cancer.
CC   Prostate specific genes P704P, P712P, P774P, P775P and B305D are located
CC   in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome
CC   region. Prostate specific antigen (PSA) P501S was located on
CC   chromosome 1. AAH84671 to AAH85143 and AAG99000 to AAG99077 represent
CC   polynucleotide and polypeptide sequences used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 3410 BP; 667 A; 1014 C; 945 G; 783 T; 1 other;

  Query Match             100.0%;  Score 3409.6;  DB 22;  Length 3410;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy            1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy           61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy          121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy          181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy          241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy          301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy          361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy       2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy       2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy       2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy       2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
```

```
Qy     2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy     2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700

Qy     2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy     2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy     2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy     2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy     2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy     3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy     3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy     3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy     3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy     3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy     3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy     3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410

RESULT 8
AAH02530
ID   AAH02530 standard; cDNA; 3410 BP.
XX
AC   AAH02530;
XX
DT   14-JUN-2001  (first entry)
XX
DE   Prostate tumour antigen determined full length cDNA sequence for L1-12.
XX
KW   Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;
KW   prostate cancer; immunogenic; cytostatic; vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200125272-A2.
XX
PD   12-APR-2001.
XX
PF   04-OCT-2000; 2000WO-US27464.
XX
PR   04-OCT-1999;   99US-0157455.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Skeiky YAW,  Reed SG,  Cheever MA;
XX
DR   WPI; 2001-245062/25.
DR   P-PSDB; AAB74800.
XX
PT   Prostate specific protein and its encoding polynucleotide, useful for
PT   the treatment and diagnosis of prostate cancer -
XX
PS   Claim 4; Page 155-156; 276pp; English.
XX
CC   The present invention describes an isolated polypeptide (I) comprising
CC   at least an immunogenic portion of a prostate tumour antigen protein or
CC   its variant. (I) have cytostatic activity and can be used in vaccine
CC   production. (I), prostate tumour antigen polynucleotides, an antigen
CC   presenting cell (APC e.g. a dendritic cell) that expresses (I), and a
CC   pharmaceutical composition containing (I) are useful for inhibiting the
CC   development of cancer in a patient. Antibodies specific for prostate
CC   specific proteins and oligonucleotides that hybridise to a
CC   polynucleotide that encodes a prostate specific protein are useful
CC   for detecting the presence or absence of a cancer or monitoring the
CC   progression the progression of a cancer, especially prostate cancer.
CC   AAH02422 to AAH2872, AAB74798 to AAB74821 and AAB74830 are sequences
CC   used in the exemplification of the present invention.
XX
SQ   Sequence 3410 BP; 667 A; 1014 C; 945 G; 783 T; 1 other;

  Query Match            100.0%;  Score 3409.6;  DB 22;  Length 3410;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy       61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy      121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy      181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy      241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy      301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy      361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy      421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy      481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCG 540

Qy        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy       2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy       2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy       2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy       2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCA 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCA 2700
```

```
Qy       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410


RESULT 9
ABL94929
ID   ABL94929 standard; cDNA; 3410 BP.
XX
AC   ABL94929;
XX
DT   19-JUL-2002  (first entry)
XX
DE   Human L1-12 cDNA sequence SEQ ID NO 110.
XX
KW   Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
KW   gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   US2002022248-A1.
XX
PD   21-FEB-2002.
XX
PF   12-JAN-2001; 2001US-0759143.
XX
PR   25-FEB-1997;   97US-0806099.
PR   01-AUG-1997;   97US-0904804.
PR   09-FEB-1998;   98US-0020956.
PR   25-FEB-1998;   98US-0030607.
PR   14-JUL-1998;   98US-0115453.
PR   23-SEP-1998;   98US-0159812.
PR   15-JAN-1999;   99US-0232149.
PR   09-APR-1999;   99US-0288946.
PR   13-JUL-1999;   99US-0352616.
PR   12-NOV-1999;   99US-0439313.
PR   18-NOV-1999;   99US-0443686.
PR   14-JAN-2000; 2000US-0483672.
PR   27-MAR-2000; 2000US-0536857.
PR   09-MAY-2000; 2000US-0568100.
PR   12-MAY-2000; 2000US-0570737.
PR   13-JUN-2000; 2000US-0593793.
PR   27-JUN-2000; 2000US-0605783.
PR   10-AUG-2000; 2000US-0636215.
PR   29-AUG-2000; 2000US-0651236.
PR   06-SEP-2000; 2000US-0657279.
PR   02-OCT-2000; 2000US-0679426.
PR   10-OCT-2000; 2000US-0685166.
XX
PA   (XUJJ/) XU J.
PA   (DILL/) DILLON D C.
PA   (MITC/) MITCHAM J L.
PA   (HARL/) HARLOCKER S L.
PA   (JIAN/) JIANG Y.
PA   (KALO/) KALOS M D.
PA   (FANG/) FANGER G R.
PA   (RETT/) RETTER M W.
PA   (STOL/) STOLK J A.
PA   (DAYC/) DAY C H.
PA   (VEDV/) VEDVICK T S.
PA   (CART/) CARTER D.
PA   (LISX/) LI S X.
PA   (WANG/) WANG A.
PA   (SKEI/) SKEIKY Y A W.
PA   (HEPL/) HEPLER W T.
PA   (HEND/) HENDERSON R A.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Kalos MD;
PI   Fanger GR,  Retter MW,  Stolk JA,  Day CH,  Vedvick TS,  Carter D;
PI   Li SX,  Wang A,  Skeiky YAW,  Hepler WT,  Henderson RA;
XX
DR   WPI; 2002-255649/30.
XX
PT   New prostate-specific polynucleotides for diagnosing and treating
PT   diseases, in particular prostate cancer, and as markers for the
PT   progression of cancer  -
XX
PS   Claim 1; SEQ ID NO 110; 87pp; English.
XX
CC   The present invention provides prostate-specific coding sequences and
CC   their encoded proteins. These can be used in the diagnosis and treatment
CC   of cancers, particularly prostate cancer. The present sequence is a cDNA
CC   described in the invention.
XX
SQ   Sequence 3410 BP; 667 A; 1014 C; 945 G; 783 T; 1 other;

  Query Match             100.0%;  Score 3409.6;  DB 24;  Length 3410;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
```

```
Qy    181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy    241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy    301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy    361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy    421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy    481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy    541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy    601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy    661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy    721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy    781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy    841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy    901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy    961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy   1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy   1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy   1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy   1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy   1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy   1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy   1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy   1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy   1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy   1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy   1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy   1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy   1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy   1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy   1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy   1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy   1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy   2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy   2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy   2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy   2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy   2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy   2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700

Qy       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
```

RESULT 10

```
ABK29017
ID   ABK29017 standard; cDNA; 3410 BP.
XX
AC   ABK29017;
XX
DT   23-APR-2002  (first entry)
XX
DE   Human breast tumour polypeptide full length cDNA clone #3.
XX
KW   Human; breast tumour polypeptide; gene; ss; breast cancer; cytostatic;
KW   immunostimulant.
XX
OS   Homo sapiens.
XX
PN   WO200198339-A2.
XX
PD   27-DEC-2001.
XX
PF   12-JUN-2001; 2001WO-US19032.
XX
PR   22-JUN-2000; 2000US-0602877.
PR   12-OCT-2000; 2000US-0687507.
PR   06-FEB-2001; 2001US-0778381.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Reed SG,  Xu J,  Dillon DC,  Retter MW,  Harlocker SL;
XX
DR   WPI; 2002-147792/19.
XX
PT   Polynucleotides encoding breast tumour polypeptides, useful for treating
PT   breast cancer or stimulating an immune response -
XX
PS   Claim 1; Page 143-144; 150pp; English.
XX
CC   The invention relates to polynucleotides encoding breast tumour
CC   polypeptides. The sequences are useful for treating cancer, preferably
CC   breast cancer, in a patient or for stimulating an immune response. The
CC   polynucleotides and polypeptides are also useful in the diagnosis and
CC   monitoring of breast cancer. A method for detecting the presence of a
CC   cancer in a patient, comprises obtaining a biological sample from the
CC   patient, contacting the biological sample with a binding agent that binds
CC   to a breast tumour polypeptide, detecting in the sample an amount of
CC   polypeptide that binds to the binding agent, and comparing the amount of
CC   polypeptide to a predetermined cut-off value, therefore determining the
CC   presence of a cancer in the patient. Sequences ABK28920-ABK29025
CC   represent cDNA clones encoding human breast tumour polypeptides of the
CC   invention.
XX
SQ   Sequence 3410 BP; 667 A; 1014 C; 945 G; 783 T; 1 other;

  Query Match             100.0%;  Score 3409.6;  DB 24;  Length 3410;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy    301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy    361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy    421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy    481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy    541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy    601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy    661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy    721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy    781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy    841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy    901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy    961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy   1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy   1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy   1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy   1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy   1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy   1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy   1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy   1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy   1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy   1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy   1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy   1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy   1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy   1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy   1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy   1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy   1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy   2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy   2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy   2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy   2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy   2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy   2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy   2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
```

```
Qy       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700

Qy       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410

RESULT 11
AAA06349
ID   AAA06349 standard; cDNA; 3410 BP.
XX
AC   AAA06349;
XX
DT   13-JUN-2000  (first entry)
XX
DE   Human immunogenic prostate tumour protein cDNA sequence SEQ ID NO:110.
XX
KW   Human; prostate cancer; diagnosis; tumour; gene therapy; detection;
KW   immunogenic; cytostatic; vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200004149-A2.
XX
PD   27-JAN-2000.
XX
PF   14-JUL-1999;   99WO-US15838.
XX
PR   14-JUL-1998;   98US-0115453.
PR   14-JUL-1998;   98US-0116134.
PR   23-SEP-1998;   98US-0159812.
PR   23-SEP-1998;   98US-0159822.
PR   15-JAN-1999;   99US-0232149.
PR   15-JAN-1999;   99US-0232880.
PR   09-APR-1999;   99US-0288946.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Dillon DC,  Harlocker SL,  Yuqiu J,  Xu J,  Mitcham JL;
XX
DR   WPI; 2000-171268/15.
XX
PT   New polypeptide useful for treating and diagnosing prostate cancer
PT   comprises an immunogenic portion of prostate tumor protein -
XX
PS   Claim 1; Page 135-136; 263pp; English.
XX
CC   The present invention describes isolated polypeptides, comprising an
CC   immunogenic portion of a prostate tumour protein (PTP). The polypeptides
CC   and polynucleotides encoding them have cytostatic activity and can be
CC   used in vaccines and in gene therapy. The polypeptides and
CC   polynucleotides encoding them, antigen presenting cells which express
CC   the polypeptides, antibodies against the polypeptides and vaccines
CC   comprising them can be used for inhibiting the development of prostate
CC   cancer in a patient. The polypeptides can be used to generate antibodies
CC   or anti-idiotypic antibodies for passive immuno therapy. A portion of
CC   the polynucleotides encoding the polypeptides can be used as a probe or
CC   to modulate the expression of the polypeptides. AAA06241 to AAA06691 and
CC   AAY82000 to AAY82020 represent sequences used in the exemplification of
CC   the present invention.
XX
SQ   Sequence 3410 BP; 667 A; 1015 C; 945 G; 782 T; 1 other;

  Query Match             99.9%;  Score 3408;  DB 21;  Length 3410;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3409;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCCG 2160

Qy       2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy       2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy       2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy       2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
```

```
Qy       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700

Qy       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
```

```
RESULT 12
AAS14962
ID   AAS14962 standard; cDNA; 3320 BP.
XX
AC   AAS14962;
XX
DT   14-FEB-2002  (first entry)
XX
DE   DNA encoding human PROST 03.
XX
KW   Human; PROST 03; metastasis; prostate cancer; tumour; immune response;
KW   cytostatic; vaccine; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             282..1943
FT                   /*tag=  a
FT                   /product= "PROST 03"
XX
PN   WO200181577-A2.
XX
PD   01-NOV-2001.
XX
PF   26-APR-2001; 2001WO-US13323.
XX
PR   27-APR-2000; 2000US-200065P.
PR   20-APR-2001; 2001US-0200065.
XX
PA   (SCHD ) SCHERING AG.
XX
PI   Lau T,  Lin RJ,  Parkes D,  Parry G,  Schneider DW,  Steinbrecher R;
PI   Van Heuit PT,  Wu J;
XX
DR   WPI; 2002-041404/05.
DR   P-PSDB; AAU10324.
XX
PT   Novel PROST 03 polypeptides and polynucleotides useful in research,
PT   diagnosis and therapeutic applications, particularly for use in cancer
PT   therapeutics -
XX
PS   Claim 6; Fig 1; 77pp; English.
XX
CC   The invention relates to an isolated PROST 03 polypeptide (I) and to
CC   the polynucleotide (II) encoding PROST 03. Fragments of (I) were used
CC   to generate antibodies (III) to PROST 03. (III) is useful for selectively
CC   destroying a cell expressing (I), and for treating a disease-state
CC   associated with expression of PROST 03 in a human patient. (III) is
CC   useful for diagnosing metastasis associated with (I), in a subject.
CC   (I) is also useful for diagnosing and treating diseases of cell
CC   proliferation such as prostate cancers. (I) is also useful for generating
CC   antibodies to PROST 03. (III) is useful in detecting the levels of PROST
CC   03 polypeptides in cells and tissues, and in targeting drugs to primary
CC   and metastatic tumours. (I) is also useful for stimulating immune
CC   response to PROST 03 containing cells. (II) is useful in diagnostic
CC   assays for detecting the levels of polynucleotides encoding PROST 03 in
CC   cells and tissues. (II) is useful as DNA probes, as targets for antisense
CC   and ribozyme therapy, and as templates for the production of antisense
CC   polynucleotides. (I) and (II) are useful in research, biological,
CC   clinical and therapeutic purposes. The present sequence represents
CC   the coding sequence of human PROST 03.
XX
SQ   Sequence 3320 BP; 585 A; 1013 C; 944 G; 778 T; 0 other;

  Query Match             96.6%;  Score 3292.4;  DB 24;  Length 3320;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3316;  Conservative    0;  Mismatches    1;  Indels    3;  Gaps    2;

Qy          3 GAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGT 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGT 60

Qy         63 GATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAA 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAA 120

Qy        123 GCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGG 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGG 180

Qy        183 CGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGA 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGA 240
```

```
Qy    243 GTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGG 302
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGG 300

Qy    303 TGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTG 362
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTG 360

Qy    363 GCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGG 422
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGG 420

Qy    423 GGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCT 482
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCT 480

Qy    483 GTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGC 542
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGC 540

Qy    543 CCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCG 602
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCG 600

Qy    603 GCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCC 662
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCC 660

Qy    663 TGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGC 722
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGC 720

Qy    723 TCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCA 782
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCA 780

Qy    783 TGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTG 842
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTG 840

Qy    843 CCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCT 902
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCT 900

Qy    903 TCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCG 962
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCG 960

Qy    963 AGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCC 1022
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 AGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCC 1020

Qy   1023 GCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCA 1082
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCA 1080

Qy   1083 TGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGA 1142
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGA 1140

Qy   1143 CCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAG 1202
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAG 1200

Qy   1203 CTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGG 1262
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGG 1260

Qy   1263 GGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGC 1322
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGC 1320

Qy   1323 AGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCG 1382
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCG 1380

Qy   1383 GTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGT 1442
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGT 1440

Qy   1443 TCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGA 1502
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 TCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGA 1500

Qy   1503 AGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCC 1562
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 AGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCC 1560

Qy   1563 TGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGG 1622
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 TGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGG 1620

Qy   1623 GTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTG 1682
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTG 1680

Qy   1683 ATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGG 1742
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 ATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGG 1740

Qy   1743 GCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCAT 1802
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCAT 1800

Qy   1803 CCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTG 1862
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 CCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTG 1860

Qy   1863 CCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCG 1922
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 CCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCG 1920

Qy   1923 ACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCA 1982
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 ACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCA 1980

Qy   1983 CTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTT 2042
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 CTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTT 2040

Qy   2043 CTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGC 2102
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 CTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGC 2100

Qy   2103 TGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGAC 2162
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 TGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGAC 2160

Qy   2163 TGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCAT 2222
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 TGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCAT 2220

Qy   2223 GCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCT 2282
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 GCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCT 2280

Qy   2283 CCTAGTTGAGACACACCTAGAGAAGGGTTTTGGGAGCTGAATAAACTCAGTCACCTGGT 2342
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CCTAGTTGAGACACACCTAGAGAAGGGTTTTGGGAGCTGAATAAACTCAGTCACCTGGT 2340

Qy   2343 TTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTT 2402
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 TTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTT 2400

Qy   2403 TCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAGA 2460
          ||||||||||||||||||||||||||||||||||||||    |  ||||||||||||||||
```

```
Db    2401 TCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGAAAGTTATTTGTAGGGGAAGA 2460

Qy    2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy    2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy    2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy    2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700

Qy    2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy    2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy    2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy    2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy    2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy    3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy    3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy    3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy    3181 GAGGTCTTATCTCTCA-GGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTAT 3239
           |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db    3181 GAGGTCTTATCTCTCAGGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTAT 3240

Qy    3240 TTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGAC 3299
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3241 TTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGAC 3300

Qy    3300 AAAATTAAAGGCTTTCTTAT 3319
           ||||||||||||||||||||
Db    3301 AAAATTAAAGGCTTTCTTAT 3320
```

```
RESULT 13
AAS64040
ID   AAS64040 standard; cDNA; 4034 BP.
XX
AC   AAS64040;
XX
DT   29-JAN-2002  (first entry)
XX
DE   Human prosate cDNA P553S splice variant #3.
XX
KW   Human; prostate cancer; ss; cytostatic; immunostimulant; tumour.
XX
OS   Homo sapiens.
XX
PN   WO200173032-A2.
XX
PD   04-OCT-2001.
XX
PF   27-MAR-2001; 2001WO-US09919.
XX
PR   27-MAR-2000; 2000US-0536857.
PR   09-MAY-2000; 2000US-0568100.
PR   12-MAY-2000; 2000US-0570737.
PR   13-JUN-2000; 2000US-0593793.
PR   27-JUN-2000; 2000US-0605783.
PR   10-AUG-2000; 2000US-0636215.
PR   29-AUG-2000; 2000US-0651236.
PR   06-SEP-2000; 2000US-0657279.
PR   02-OCT-2000; 2000US-0679426.
PR   10-OCT-2000; 2000US-0685166.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Kalos MD;
PI   Fanger GR,  Retter MW,  Stolk JA,  Day CH,  Vedvick TS,  Carter D;
PI   Li SX,  Wang A,  Skeiky YAW,  Hepler WT,  Henderson RA;
XX
DR   WPI; 2001-639232/73.
XX
PT   New human prostate-specific polypeptides and polynucleotides useful for
PT   the diagnosis and treatment of cancer, especially prostate cancer -
XX
PS   Claim 1; Page 483-484; 579pp; English.
XX
CC   The invention relates to isolated prostate-specific
CC   polynucleotides, polypeptides, fusion proteins of the polypeptides,
CC   antibodies raised against the polypeptides (or antigenic epitopes
CC   derived from them) and antigen-presenting cells expressing the
CC   polypeptides. The antibodies are useful for detecting the presence of
CC   cancer, especially prostate cancer. The polypeptides, polynucleotides and
CC   the antigen-presenting cells are useful for stimulating and/or expanding
CC   T cells specific for a tumour protein, and for inhibiting the development
CC   of cancer especially prostate cancer. Compositions comprising the
CC   polynucleotide and/or polypeptide are useful for stimulating an immune
CC   response, and for treating cancer. The oligonucleotide is useful for
CC   detecting cancer. The present sequence is a prostate specific
CC   polynucleotide of the invention.
XX
SQ   Sequence 4034 BP; 721 A; 1226 C; 1141 G; 946 T; 0 other;

  Query Match             75.8%;  Score 2585.4;  DB 22;  Length 4034;
  Best Local Similarity   82.4%;  Pred. No. 0;
  Matches 3325;  Conservative    0;  Mismatches    1;  Indels  709;  Gaps    3;

Qy          4 AACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTG 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTG 60

Qy         64 ATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAG 120

Qy        124 CTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGC 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGC 180

Qy        184 GGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAG 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAG 240

Qy        244 TGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGT 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGT 300
```

```
Qy        304 GAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGG 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGG 360

Qy        364 CCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGG 420

Qy        424 GGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTG 480

Qy        484 TGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC 540

Qy        544 CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGG 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGG 600

Qy        604 CTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCT 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCT 660

Qy        664 GGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT 720

Qy        724 CTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCAT 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCAT 780

Qy        784 GATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGC 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGC 840

Qy        844 CCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTT 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTT 900

Qy        904 CCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGA 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGA 960

Qy        964 GCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCG 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCG 1020

Qy       1024 CTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCAT 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCAT 1080

Qy       1084 GCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGAC 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGAC 1140

Qy       1144 CTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGC 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGC 1200

Qy       1204 TGAGCCGGGCACCGAGGCCCGGAGACACTATGATGA------------------------ 1239
              ||||||||||||||||||||||||||||||||||||
Db       1201 TGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGTAAGGCCTTGGCAGCCAGCAG 1260

Qy       1240 ------------------------------------------------------------ 1239

Db       1261 AGGCTGGTGTGGAGCCGCCCACCAGAGACGACACTCGGGGCTGTGTCTGGGCTGGTGCC 1320

Qy       1240 ------------------------------------------------------------ 1239

Db       1321 TCTCCATCCTGGCCCCGACTTCTCTGTCAGGAAAGTGGGGATGGACCCCATCTGCATACA 1380

Qy       1240 ------------------------------------------------------------ 1239

Db       1381 CGGCTTCTCATGGGTGTGGAACATCTCTGCTTGCGGTTTCAGGAAGGCCTCTGGCTGCTC 1440

Qy       1240 ------------------------------------------------------------ 1239

Db       1441 TAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGACAGAAGGAAAGGCGGAGCTTATTCAA 1500

Qy       1240 ------------------------------------------------------------ 1239

Db       1501 AGTCTAGAGGGAGTGGAGGAGTTAAGGCTGGATTTCAGATCTGCCTGGTTCCAGCCGCAG 1560

Qy       1240 ------------------------------------------------------------ 1239

Db       1561 TGTGCCCTCTGCTCCCCCAACGACTTTCCAAATAATCTCACCAGCGCCTTCCAGCTCAGG 1620

Qy       1240 ------------------------------------------------------------ 1239

Db       1621 CGTCCTAGAAGCGTCTTGAAGCCTATGGCCAGCTGTCTTTGTGTTCCCTCTCACCCGCCT 1680

Qy       1240 ------------------------------------------------------------ 1239

Db       1681 GTCCTCACAGCTGAGACTCCCAGGAAACCTTCAGACTACCTTCCTCTGCCTTCAGCAAGG 1740

Qy       1240 ------------------------------------------------------------ 1239

Db       1741 GGCGTTGCCCACATTCTCTGAGGGTCAGTGGAAGAACCTAGACTCCCATTGCTAGAGGTA 1800

Qy       1240 ------------------------------------------------------------ 1239

Db       1801 GAAAGGGGAAGGGTGCTGGGGAGCAGGGCTGGTCCACAGCAGGTCTCGTGCAGCAGGTAC 1860

Qy       1240 ------------------------------------------------------------ 1239

Db       1861 CTGTGGTTCCGCCTTCTCATCTCCCTGAGACTGCTCCGACCCTTCCCTCCCAGGCTCTGT 1920

Qy       1240 ----------------------AGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGT 1277
                                    ||||||||||||||||||||||||||||||||||||||
Db       1921 CTGATGGCCCCTCTCCCTCTGCAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGT 1980

Qy       1278 GCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTC 1337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTC 2040

Qy       1338 GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT 1397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT 2100

Qy       1398 CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCC 1457
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCC 2160

Qy       1458 TGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGC 1517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 TGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGC 2220

Qy       1518 CCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCC 1577
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCC 2280

Qy       1578 TGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTG 1637
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 TGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTG 2340

Qy       1638 GCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTG 1697
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTG 2400

Qy       1698 TGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACC 1757
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACC 2460

Qy       1758 TCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCT 1817
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 TCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCT 2520

Qy       1818 CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTC 1877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTC 2580

Qy       1878 TGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT 1937
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 TGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT 2640

Qy       1938 CAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCC 1997
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCC 2700

Qy       1998 CCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAG 2057
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 CCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAG 2760

Qy       2058 TAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2117
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2820

Qy       2118 TGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAG 2177
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAG 2880

Qy       2178 GGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGG 2237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGG 2940

Qy       2238 GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACA 2297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACA 3000

Qy       2298 CCTAGAGAAGGGTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGC 2357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCTAGAGAAGGGTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGC 3060

Qy       2358 CCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACAC 2417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACAC 3120

Qy       2418 TCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAAC 2475
              ||||||||||||||||||||||||   | ||||||||||||||||||||||||||||||||
Db       3121 TCCTCCATGGGATTTGAACATATGAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAAC 3180

Qy       2476 ACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCT 2535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 ACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCT 3240

Qy       2536 TACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCAT 2595
              |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db       3241 TACCTTTTATCAGGATGT-GCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCAT 3299

Qy       2596 TTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGT 2655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3300 TTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGT 3359

Qy       2656 GTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAG 2715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3360 GTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAG 3419

Qy       2716 CTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAAT 2775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3420 CTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAAT 3479

Qy       2776 GCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTAC 2835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3480 GCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTAC 3539

Qy       2836 CCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT 2895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3540 CCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT 3599

Qy       2896 AACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTC 2955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3600 AACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTC 3659

Qy       2956 TCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTA 3015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3660 TCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTA 3719

Qy       3016 CCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCA 3075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3720 CCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCA 3779

Qy       3076 CAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGG 3135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3780 CAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGG 3839

Qy       3136 GAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTC 3195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3840 GAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTC 3899

Qy       3196 AGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATAT 3255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3900 AGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATAT 3959

Qy       3256 TTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTC 3315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3960 TTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTC 4019

Qy       3316 TTATATGTTTAAAAA 3330
              |||||||||||||||
Db       4020 TTATATGTTTAAAAA 4034
```

```
RESULT 14
AAH93868
ID   AAH93868 standard; cDNA; 4034 BP.
XX
AC   AAH93868;
XX
DT   04-OCT-2001  (first entry)
XX
DE   P553S cDNA splice variant P553S-10.
XX
KW   Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW   cytostatic; gene therapy; metastasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200151633-A2.
XX
PD   19-JUL-2001.
XX
PF   16-JAN-2001; 2001WO-US01574.
XX
PR   14-JAN-2000; 2000US-0483672.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Reed SG;
PI   Kalos MD,  Fanger GR,  Day CH,  Retter MW,  Stolk JA,  Skeiky YAW;
PI   Wang A,  Meagher MJ;
XX
DR   WPI; 2001-425873/45.
XX
PT   New polynucleotide encoding a prostate-specific protein, for
PT   diagnosing, monitoring and treating prostate cancer in a patient and
PT   for use in vaccines -
XX
PS   Claim 1; Page 460-461; 543pp; English.
XX
CC   The present invention describes polynucleotide sequences (I) which encode
```

```
CC   prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC   and can be used in vaccine production and gene therapy. (I), (II),
CC   antibodies to (II), fusion proteins comprising (II), and isolated
CC   T cells prepared using (I) or (II) are used treat cancer in a patient.
CC   (I) and the antibodies are also used in the detection of cancer in a
CC   patient. The cancer that is diagnosed or treated is particularly
CC   prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
CC   (I) can be used for monitoring the progression of cancer in a patient.
CC   (I) and (II) can also be used to improve diagnostic and therapeutic
CC   methods for prostate cancer. They can indicate the level of metastasis
CC   as well as the prostate volume. AAH93357 to AAH93944 and AAM01115 to
CC   AAM01318 represent polynucleotide and amino acid sequences used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 4034 BP; 721 A; 1226 C; 1141 G; 946 T; 0 other;

  Query Match            75.8%;  Score 2585.4;  DB 22;  Length 4034;
  Best Local Similarity  82.4%;  Pred. No. 0;
  Matches 3325;  Conservative    0;  Mismatches    1;  Indels  709;  Gaps    3;

Qy          4 AACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTG 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTG 60

Qy         64 ATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAG 120

Qy        124 CTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGC 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGC 180

Qy        184 GGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAG 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAG 240

Qy        244 TGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGT 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGT 300

Qy        304 GAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGG 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGG 360

Qy        364 CCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGG 420

Qy        424 GGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTG 480

Qy        484 TGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC 540

Qy        544 CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGG 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGG 600

Qy        604 CTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCT 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCT 660

Qy        664 GGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT 720

Qy        724 CTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCAT 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCAT 780

Qy        784 GATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGC 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGC 840

Qy        844 CCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTT 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTT 900

Qy        904 CCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGA 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGA 960

Qy        964 GCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCG 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCG 1020

Qy       1024 CTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCAT 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCAT 1080

Qy       1084 GCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGAC 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGAC 1140

Qy       1144 CTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGC 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGC 1200

Qy       1204 TGAGCCGGGCACCGAGGCCCGGAGACACTATGATGA------------------------ 1239
              ||||||||||||||||||||||||||||||||||||
Db       1201 TGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGTAAGGCCTTGGCAGCCAGCAG 1260

Qy       1240 ------------------------------------------------------------ 1239

Db       1261 AGGCTGGTGTGGGAGCCGCCCACCAGAGACGACACTCGGGGCTGTGTCTGGGCTGGTGCC 1320

Qy       1240 ------------------------------------------------------------ 1239

Db       1321 TCTCCATCCTGGCCCCGACTTCTCTGTCAGGAAAGTGGGGATGGACCCCATCTGCATACA 1380

Qy       1240 ------------------------------------------------------------ 1239

Db       1381 CGGCTTCTCATGGGTGTGGAACATCTCTGCTTGCGGTTTCAGGAAGGCCTCTGGCTGCTC 1440

Qy       1240 ------------------------------------------------------------ 1239

Db       1441 TAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGACAGAAGGAAAGGCGGAGCTTATTCAA 1500

Qy       1240 ------------------------------------------------------------ 1239

Db       1501 AGTCTAGAGGGAGTGGAGGAGTTAAGGCTGGATTTCAGATCTGCCTGGTTCCAGCCGCAG 1560

Qy       1240 ------------------------------------------------------------ 1239

Db       1561 TGTGCCCTCTGCTCCCCCAACGACTTTCCAAATAATCTCACCAGCGCCTTCCAGCTCAGG 1620

Qy       1240 ------------------------------------------------------------ 1239

Db       1621 CGTCCTAGAAGCGTCTTGAAGCCTATGGCCAGCTGTCTTTGTGTTCCCTCTCACCCGCCT 1680

Qy       1240 ------------------------------------------------------------ 1239

Db       1681 GTCCTCACAGCTGAGACTCCCAGGAAACCTTCAGACTACCTTCCTCTGCCTTCAGCAAGG 1740

Qy       1240 ------------------------------------------------------------ 1239

Db       1741 GGCGTTGCCCACATTCTCTGAGGGTCAGTGGAAGAACCTAGACTCCCATTGCTAGAGGTA 1800

Qy       1240 ------------------------------------------------------------ 1239

Db       1801 GAAAGGGGAAGGGTGCTGGGGAGCAGGGCTGGTCCACAGCAGGTCTCGTGCAGCAGGTAC 1860

Qy       1240 ------------------------------------------------------------ 1239
```

```
Db   1861 CTGTGGTTCCGCCTTCTCATCTCCCTGAGACTGCTCCGACCCTTCCCTCCCAGGCTCTGT 1920

Qy   1240 ----------------------AGGCGTTCGGATGGCAGCCTGGGGCTGTTCCTGCAGT 1277
                                ||||||||||||||||||||||||||||||||||||||
Db   1921 CTGATGGCCCCTCTCCCTCTGCAGGCGTTCGGATGGCAGCCTGGGGCTGTTCCTGCAGT 1980

Qy   1278 GCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTC 1337
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 GCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTC 2040

Qy   1338 GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT 1397
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT 2100

Qy   1398 CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCC 1457
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCC 2160

Qy   1458 TGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGC 1517
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 TGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGC 2220

Qy   1518 CCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCC 1577
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 CCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCC 2280

Qy   1578 TGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTG 1637
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 TGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTG 2340

Qy   1638 GCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTG 1697
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTG 2400

Qy   1698 TGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACC 1757
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACC 2460

Qy   1758 TCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCT 1817
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 TCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCT 2520

Qy   1818 CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTC 1877
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTC 2580

Qy   1878 TGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT 1937
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 TGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT 2640

Qy   1938 CAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCC 1997
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 CAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCC 2700

Qy   1998 CCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAG 2057
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 CCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAG 2760

Qy   2058 TAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2117
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 TAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2820

Qy   2118 TGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAG 2177
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 TGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAG 2880

Qy   2178 GGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGG 2237
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 GGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGG 2940

Qy   2238 GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACA 2297
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACA 3000

Qy   2298 CCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGC 2357
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 CCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGC 3060

Qy   2358 CCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACAC 2417
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 CCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACAC 3120

Qy   2418 TCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAAC 2475
          ||||||||||||||||||||||||  | ||||||||||||||||||||||||||||||||
Db   3121 TCCTCCATGGGATTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAAC 3180

Qy   2476 ACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCT 2535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181 ACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCT 3240

Qy   2536 TACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCAT 2595
          |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   3241 TACCTTTTATCAGGATGT-GCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCAT 3299

Qy   2596 TTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGT 2655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3300 TTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGT 3359

Qy   2656 GTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAG 2715
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3360 GTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAG 3419

Qy   2716 CTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAAT 2775
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3420 CTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAAT 3479

Qy   2776 GCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTAC 2835
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3480 GCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTAC 3539

Qy   2836 CCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT 2895
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3540 CCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT 3599

Qy   2896 AACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTC 2955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3600 AACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTC 3659

Qy   2956 TCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTA 3015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3660 TCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTA 3719

Qy   3016 CCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCA 3075
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3720 CCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCA 3779

Qy   3076 CAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGG 3135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3780 CAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGG 3839

Qy   3136 GAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTC 3195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3840 GAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTC 3899

Qy   3196 AGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATAT 3255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3900 AGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATAT 3959

Qy   3256 TTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTC 3315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3960 TTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTC 4019

Qy   3316 TTATATGTTTAAAAA 3330
          |||||||||||||||
Db   4020 TTATATGTTTAAAAA 4034
```

```
RESULT 15
ABL95411
ID   ABL95411 standard; cDNA; 4034 BP.
XX
AC   ABL95411;
XX
DT   19-JUL-2002  (first entry)
XX
DE   Human P553S splice variant SEQ ID NO 704.
XX
KW   Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
KW   gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   US2002022248-A1.
XX
PD   21-FEB-2002.
XX
PF   12-JAN-2001; 2001US-0759143.
XX
PR   25-FEB-1997;    97US-0806099.
PR   01-AUG-1997;    97US-0904804.
PR   09-FEB-1998;    98US-0020956.
PR   25-FEB-1998;    98US-0030607.
PR   14-JUL-1998;    98US-0115453.
PR   23-SEP-1998;    98US-0159812.
PR   15-JAN-1999;    99US-0232149.
PR   09-APR-1999;    99US-0288946.
PR   13-JUL-1999;    99US-0352616.
PR   12-NOV-1999;    99US-0439313.
PR   18-NOV-1999;    99US-0443686.
PR   14-JAN-2000;  2000US-0483672.
PR   27-MAR-2000;  2000US-0536857.
PR   09-MAY-2000;  2000US-0568100.
PR   12-MAY-2000;  2000US-0570737.
PR   13-JUN-2000;  2000US-0593793.
PR   27-JUN-2000;  2000US-0605783.
PR   10-AUG-2000;  2000US-0636215.
PR   29-AUG-2000;  2000US-0651236.
PR   06-SEP-2000;  2000US-0657279.
PR   02-OCT-2000;  2000US-0679426.
PR   10-OCT-2000;  2000US-0685166.
XX
PA   (XUJJ/) XU J.
PA   (DILL/) DILLON D C.
PA   (MITC/) MITCHAM J L.
PA   (HARL/) HARLOCKER S L.
PA   (JIAN/) JIANG Y.
PA   (KALO/) KALOS M D.
PA   (FANG/) FANGER G R.
PA   (RETT/) RETTER M W.
PA   (STOL/) STOLK J A.
PA   (DAYC/) DAY C H.
PA   (VEDV/) VEDVICK T S.
PA   (CART/) CARTER D.
PA   (LISX/) LI S X.
PA   (WANG/) WANG A.
PA   (SKEI/) SKEIKY Y A W.
PA   (HEPL/) HEPLER W T.
PA   (HEND/) HENDERSON R A.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Kalos MD;
PI   Fanger GR,  Retter MW,  Stolk JA,  Day CH,  Vedvick TS,  Carter D;
PI   Li SX,  Wang A,  Skeiky YAW,  Hepler WT,  Henderson RA;
XX
DR   WPI; 2002-255649/30.
XX
PT   New prostate-specific polynucleotides for diagnosing and treating
PT   diseases, in particular prostate cancer, and as markers for the
PT   progression of cancer  -
XX
PS   Claim 1; SEQ ID NO 704; 87pp; English.
XX
CC   The present invention provides prostate-specific coding sequences and
CC   their encoded proteins. These can be used in the diagnosis and treatment
CC   of cancers, particularly prostate cancer. The present sequence is a cDNA
CC   described in the invention.
XX
SQ   Sequence 4034 BP; 721 A; 1226 C; 1141 G; 946 T; 0 other;

  Query Match             75.8%;  Score 2585.4;  DB 24;  Length 4034;
  Best Local Similarity   82.4%;  Pred. No. 0;
  Matches 3325;  Conservative    0;  Mismatches    1;  Indels  709;  Gaps    3;

Qy          4 AACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTG 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTG 60

Qy         64 ATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAG 120

Qy        124 CTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGC 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGC 180

Qy        184 GGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAG 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAG 240

Qy        244 TGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGT 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGT 300

Qy        304 GAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGG 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGG 360

Qy        364 CCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGG 420

Qy        424 GGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTG 480

Qy        484 TGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC 540

Qy        544 CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGG 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGG 600

Qy        604 CTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCT 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCT 660

Qy        664 GGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT 720

Qy        724 CTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCAT 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCAT 780

Qy        784 GATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGC 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGC 840

Qy        844 CCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTT 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTT 900
```

```
Qy        904 CCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGA 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGA 960

Qy        964 GCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCG 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCG 1020

Qy       1024 CTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCAT 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCAT 1080

Qy       1084 GCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGAC 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGAC 1140

Qy       1144 CTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGC 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGC 1200

Qy       1204 TGAGCCGGGCACCGAGGCCCGGAGACACTATGATGA------------------------ 1239
              ||||||||||||||||||||||||||||||||||||
Db       1201 TGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGTAAGGCCTTGGCAGCCAGCAG 1260

Qy       1240 ------------------------------------------------------------ 1239

Db       1261 AGGCTGGTGTGGGAGCCGCCCACCAGAGACGACACTCGGGGCTGTGTCTGGGCTGGTGCC 1320

Qy       1240 ------------------------------------------------------------ 1239

Db       1321 TCTCCATCCTGGCCCCGACTTCTCTGTCAGGAAAGTGGGGATGGACCCCATCTGCATACA 1380

Qy       1240 ------------------------------------------------------------ 1239

Db       1381 CGGCTTCTCATGGGTGTGGAACATCTCTGCTTGCGGTTTCAGGAAGGCCTCTGGCTGCTC 1440

Qy       1240 ------------------------------------------------------------ 1239

Db       1441 TAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGACAGAAGGAAAGGCGGAGCTTATTCAA 1500

Qy       1240 ------------------------------------------------------------ 1239

Db       1501 AGTCTAGAGGGAGTGGAGGAGTTAAGGCTGGATTTCAGATCTGCCTGGTTCCAGCCGCAG 1560

Qy       1240 ------------------------------------------------------------ 1239

Db       1561 TGTGCCCTCTGCTCCCCCAACGACTTTCCAAATAATCTCACCAGCGCCTTCCAGCTCAGG 1620

Qy       1240 ------------------------------------------------------------ 1239

Db       1621 CGTCCTAGAAGCGTCTTGAAGCCTATGGCCAGCTGTCTTTGTGTTCCCTCTCACCCGCCT 1680

Qy       1240 ------------------------------------------------------------ 1239

Db       1681 GTCCTCACAGCTGAGACTCCCAGGAAACCTTCAGACTACCTTCCTCTGCCTTCAGCAAGG 1740

Qy       1240 ------------------------------------------------------------ 1239

Db       1741 GGCGTTGCCCACATTCTCTGAGGGTCAGTGGAAGAACCTAGACTCCCATTGCTAGAGGTA 1800

Qy       1240 ------------------------------------------------------------ 1239

Db       1801 GAAAGGGGAAGGGTGCTGGGGAGCAGGGCTGGTCCACAGCAGGTCTCGTGCAGCAGGTAC 1860

Qy       1240 ------------------------------------------------------------ 1239

Db       1861 CTGTGGTTCCGCCTTCTCATCTCCCTGAGACTGCTCCGACCCTTCCCTCCCAGGCTCTGT 1920

Qy       1240 ----------------------AGGCGTTCGGATGGCAGCCTGGGCTGTTCCTGCAGT 1277
                                    ||||||||||||||||||||||||||||||||||||||
Db       1921 CTGATGGCCCCTCTCCCTCTGCAGGCGTTCGGATGGCAGCCTGGGCTGTTCCTGCAGT 1980

Qy       1278 GCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTC 1337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTC 2040

Qy       1338 GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT 1397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT 2100

Qy       1398 CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCC 1457
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCC 2160

Qy       1458 TGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGC 1517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 TGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGC 2220

Qy       1518 CCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCC 1577
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCC 2280

Qy       1578 TGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTG 1637
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 TGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTG 2340

Qy       1638 GCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTG 1697
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTG 2400

Qy       1698 TGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACC 1757
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACC 2460

Qy       1758 TCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCT 1817
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 TCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCT 2520

Qy       1818 CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTC 1877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTC 2580

Qy       1878 TGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT 1937
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 TGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT 2640

Qy       1938 CAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCC 1997
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCC 2700

Qy       1998 CCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAG 2057
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 CCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAG 2760

Qy       2058 TAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2117
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2820

Qy       2118 TGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAG 2177
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAG 2880

Qy       2178 GGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGG 2237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGG 2940

Qy       2238 GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACA 2297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACA 3000

Qy       2298 CCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGC 2357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGC 3060

Qy       2358 CCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACAC 2417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        3061 CCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACAC 3120

Qy        2418 TCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAAC 2475
               ||||||||||||||||||||||||  | |||||||||||||||||||||||||||||||||
Db        3121 TCCTCCATGGGATTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAAC 3180

Qy        2476 ACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCT 2535
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3181 ACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCT 3240

Qy        2536 TACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCAT 2595
               |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        3241 TACCTTTTATCAGGATGT-GCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCAT 3299

Qy        2596 TTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGT 2655
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3300 TTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGT 3359

Qy        2656 GTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAG 2715
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3360 GTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAG 3419

Qy        2716 CTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAAT 2775
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3420 CTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAAT 3479

Qy        2776 GCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTAC 2835
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3480 GCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTAC 3539

Qy        2836 CCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT 2895
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3540 CCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT 3599

Qy        2896 AACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTC 2955
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3600 AACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTC 3659

Qy        2956 TCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTA 3015
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3660 TCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTA 3719

Qy        3016 CCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCA 3075
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3720 CCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCA 3779

Qy        3076 CAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGG 3135
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3780 CAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGG 3839

Qy        3136 GAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTC 3195
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3840 GAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTC 3899

Qy        3196 AGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATAT 3255
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3900 AGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATAT 3959

Qy        3256 TTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTC 3315
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3960 TTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTC 4019

Qy        3316 TTATATGTTTAAAAA 3330
               |||||||||||||||
Db        4020 TTATATGTTTAAAAA 4034
```

RESULT 16
ABK92217
```
ID   ABK92217 standard; DNA; 2582 BP.
XX
AC   ABK92217;
XX
DT   15-AUG-2002  (first entry)
XX
DE   Prostate cancer-associated DNA sequence #103.
XX
KW   Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;
KW   gene therapy; gene; ds.
XX
OS   Mammalia.
XX
PN   WO200230268-A2.
XX
PD   18-APR-2002.
XX
PF   12-OCT-2001; 2001WO-US32045.
XX
PR   13-OCT-2000; 2000US-0687576.
PR   08-DEC-2000; 2000US-0733288.
PR   08-DEC-2000; 2000US-0733742.
PR   24-JAN-2001; 2001US-263957P.
PR   16-MAR-2001; 2001US-276791P.
PR   16-MAR-2001; 2001US-276888P.
PR   06-APR-2001; 2001US-281922P.
PR   24-APR-2001; 2001US-286214P.
PR   30-APR-2001; 2001US-0847046.
PR   04-MAY-2001; 2001US-288589P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Gish KC,  Mack DH,  Wilson KE,  Afar D,  Hevezi P;
XX
DR   WPI; 2002-471335/50.
DR   P-PSDB; ABG61900.
XX
PT   Detecting a prostate cancer-associated transcript in a cell in a
PT   patient, useful for diagnosing prostate cancer (PC) or screening
PT   modulators of PC, by determining if prostate cancer-associated genes
PT   are expressed in a prostate tissue  -
XX
PS   Claim 22; Page 386; 436pp; English.
XX
CC   The present invention relates to methods of detecting a prostate
CC   cancer-associated transcript in a cell from a patient. The method
CC   comprises contacting a biological sample from the patient with
CC   prostate cancer-associated polynucleotides (designated PC genes) that
CC   selectively hybridise to a sequence that is at least 80% identical
CC   to them. The prostate cancer-associated polynucleotide sequences
CC   are differentially,expressed in prostate tumour tissue or in
CC   prostate cancer and are derived from the tissues of various
CC   organisms such as humans or other mammals (e.g. mice, sheep and dogs).
CC   The methods of the invention are useful for diagnosing and treating
CC   prostate cancer in mammals. The prostate cancer-associated genes are
CC   useful for diagnosing or treating prostate cancer, as well as for
CC   identifying modulators of prostate cancer or agents that inhibit
CC   prostate cancer. The nucleic acid sequences are particularly useful
CC   in gene therapy, as a vaccine or in antisense applications.
CC   ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
CC   sequences.
XX
SQ   Sequence 2582 BP; 407 A; 822 C; 785 G; 566 T; 2 other;

  Query Match             72.9%;  Score 2486.4;  DB 24;  Length 2582;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 2499;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    1;

Qy          2 GGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAG 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         28 GGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAG 87

Qy         62 TGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGA 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 TGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGA 147

Qy        122 AGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTG 181
```

```
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 AGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTG 207

Qy        182 GCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGG 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 GCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGG 267

Qy        242 AGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGG 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 AGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGG 327

Qy        302 GTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTT 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 GTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTT 387

Qy        362 GGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTG 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 GGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTG 447

Qy        422 GGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTC 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 GGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTC 507

Qy        482 TGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGG 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 TGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGG 567

Qy        542 CCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCC 601
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 CCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCC 627

Qy        602 GGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATC 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 GGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATC 687

Qy        662 CTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTG 721
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 CTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTG 747

Qy        722 CTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTC 781
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTC 807

Qy        782 ATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGT 841
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 ATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGT 867

Qy        842 GCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATC 901
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 GCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATC 927

Qy        902 TTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACC 961
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACC 987

Qy        962 GAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCC 1021
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 GAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCC 1047

Qy       1022 CGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGC 1081
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 CGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGC 1107

Qy       1082 ATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATG 1141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 ATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATG 1167

Qy       1142 ACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGA 1201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 ACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGA 1227

Qy       1202 GCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTG 1261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1228 GCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTG 1287

Qy       1262 GGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTG 1321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 GGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTG 1347

Qy       1322 CAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCC 1381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 CAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCC 1407

Qy       1382 GGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGG 1441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 GGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGG 1467

Qy       1442 TTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAG 1501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 TTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAG 1527

Qy       1502 AAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGC 1561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1528 AAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGC 1587

Qy       1562 CTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTG 1621
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1588 CTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTG 1647

Qy       1622 GGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGT 1681
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1648 GGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGT 1707

Qy       1682 GATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGG 1741
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1708 GATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGG 1767

Qy       1742 GGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCA 1801
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1768 GGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCA 1827

Qy       1802 TCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCT 1861
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1828 TCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCT 1887

Qy       1862 GCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGC 1921
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1888 GCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGC 1947

Qy       1922 GACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTC 1981
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1948 GACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTC 2007

Qy       1982 ACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTT 2041
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2008 ACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTT 2067

Qy       2042 TCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAG 2101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2068 TCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAG 2127

Qy       2102 CTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGA 2161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2128 CTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGA 2187

Qy       2162 CTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCA 2221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2188 CTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCA 2247

Qy       2222 TGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCC 2281
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2248 TGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCC 2307

Qy       2282 TCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGG 2341
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2308 TCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGG 2367
```

```
Qy       2342 TTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGT 2401
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2368 TTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGT 2427

Qy       2402 TTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAG 2459
              ||||||||||||||||||||||||||||||||||||||||   | ||||||||||||||||
Db       2428 TTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGAAAGTTATTTGTAGGGGAAG 2487

Qy       2460 AGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCC 2501
              ||||||||||||||||||||||||||||||||||||||||||
Db       2488 AGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCC 2529
```

RESULT 17
AAS64039
ID   AAS64039 standard; cDNA; 2904 BP.
XX
AC   AAS64039;
XX
DT   29-JAN-2002 (first entry)
XX
DE   Human prosate cDNA P553S splice variant #2.
XX
KW   Human; prostate cancer; ss; cytostatic; immunostimulant; tumour.
XX
OS   Homo sapiens.
XX
PN   WO200173032-A2.
XX
PD   04-OCT-2001.
XX
PF   27-MAR-2001; 2001WO-US09919.
XX
PR   27-MAR-2000; 2000US-0536857.
PR   09-MAY-2000; 2000US-0568100.
PR   12-MAY-2000; 2000US-0570737.
PR   13-JUN-2000; 2000US-0593793.
PR   27-JUN-2000; 2000US-0605783.
PR   10-AUG-2000; 2000US-0636215.
PR   29-AUG-2000; 2000US-0651236.
PR   06-SEP-2000; 2000US-0657279.
PR   02-OCT-2000; 2000US-0679426.
PR   10-OCT-2000; 2000US-0685166.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI   Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
PI   Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
XX
DR   WPI; 2001-639232/73.
XX
PT   New human prostate-specific polypeptides and polynucleotides useful for
PT   the diagnosis and treatment of cancer, especially prostate cancer -
XX
PS   Claim 1; Page 482-483; 579pp; English.
XX
CC   The invention relates to isolated prostate-specific
CC   polynucleotides, polypeptides, fusion proteins of the polypeptides,
CC   antibodies raised against the polypeptides (or antigenic epitopes
CC   derived from them) and antigen-presenting cells expressing the
CC   polypeptides. The antibodies are useful for detecting the presence of
CC   cancer, especially prostate cancer. The polypeptides, polynucleotides and
CC   the antigen-presenting cells are useful for stimulating and/or expanding
CC   T cells specific for a tumour protein, and for inhibiting the development
CC   of cancer especially prostate cancer. Compositions comprising the
CC   polynucleotide and/or polypeptide are useful for stimulating an immune
CC   response, and for treating cancer. The oligonucleotide is useful for
CC   detecting cancer. The present sequence is a prostate specific
CC   polynucleotide of the invention.
XX
SQ   Sequence 2904 BP; 542 A; 875 C; 773 G; 714 T; 0 other;

```
  Query Match             64.4%;  Score 2196.4;  DB 22;  Length 2904;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches 2560;  Conservative    0;  Mismatches    1;  Indels  343;  Gaps    2;

Qy        770 GTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC 829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC 60

Qy        830 TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTG 889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTG 120

Qy        890 CTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCG 949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCG 180

Qy        950 CTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGT 1009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGT 240

Qy       1010 CCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAG 1069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAG 300

Qy       1070 CTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGG 1129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGG 360

Qy       1130 ATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAG 1189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAG 420

Qy       1190 GGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAA--------- 1240
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGAAGGCCT 480

Qy       1241 ------------------------------------------------------------ 1240

Db        481 CTGGCTGCTCTAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGACAGAAGGAAAGGCGGA 540

Qy       1241 ------------------------------------------------------------ 1240

Db        541 GCTTATTCAAAGTCTAGAGGGAGTGGAGGAGTTAAGGCTGGATTTCAGATCTGCCTGGTT 600

Qy       1241 ------------------------------------------------------------ 1240

Db        601 CCAGCCGCAGTGTGCCCTCTGCTCCCCCAACGACTTTCCAAATAATCTCACCAGCGCCTT 660

Qy       1241 ------------------------------------------------------------ 1240

Db        661 CCAGCTCAGGCGTCCTAGAAGCGTCTTGAAGCCTATGGCCAGCTGTCTTTGTGTTCCCTC 720

Qy       1241 ------------------------------------------------------------ 1240

Db        721 TCACCCGCCTGTCCTCACAGCTGAGACTCCCAGGAAACCTTCAGACTACCTTCCTCTGCC 780

Qy       1241 --------------------------------GGCGTTCGGATGGGCAGCCTGGGGCTGT 1268
                                              ||||||||||||||||||||||||||||
Db        781 TTCAGCAAGGGGCGTTGCCCACATTCTCTGAGGGCGTTCGGATGGGCAGCCTGGGGCTGT 840

Qy       1269 TCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGAT 1328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGAT 900

Qy       1329 TCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCA 1388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCA 960

Qy       1389 CATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCT 1448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCT 1020
```

```
Qy       1449 TCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGG 1508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGG 1080

Qy       1509 TGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGA 1568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGA 1140

Qy       1569 CCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTG 1628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTG 1200

Qy       1629 GAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCT 1688
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCT 1260

Qy       1689 CCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCT 1748
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCT 1320

Qy       1749 GCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGT 1808
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGT 1380

Qy       1809 TTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAG 1868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAG 1440

Qy       1869 GCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGG 1928
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGG 1500

Qy       1929 CCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGT 1988
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGT 1560

Qy       1989 CCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTG 2048
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTG 1620

Qy       2049 CTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACA 2108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACA 1680

Qy       2109 GCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGG 2168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGG 1740

Qy       2169 CCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTG 2228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTG 1800

Qy       2229 GAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGT 2288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGT 1860

Qy       2289 TGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCA 2348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCA 1920

Qy       2349 TCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGG 2408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGG 1980

Qy       2409 ATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTG 2466
              |||||||||||||||||||||||||||||||||  | |||||||||||||||||||||||
Db       1981 ATGAAACACTCCTCCATGGGATTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTG 2040

Qy       2467 AGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCA 2526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 AGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCA 2100

Qy       2527 CCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGA 2586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGA 2160

Qy       2587 CACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAG 2646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 CACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAG 2220

Qy       2647 CTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCC 2706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCC 2280

Qy       2707 CTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCC 2766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCC 2340

Qy       2767 CCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAA 2826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 CCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAA 2400

Qy       2827 TGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAAC 2886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAAC 2460

Qy       2887 GGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCC 2946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCC 2520

Qy       2947 TCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAAC 3006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAAC 2580

Qy       3007 TTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGA 3066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 TTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGA 2640

Qy       3067 CCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCT 3126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCT 2700

Qy       3127 GTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTC 3186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTC 2760

Qy       3187 TTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGG 3246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGG 2820

Qy       3247 GGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTA 3306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 GGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTA 2880

Qy       3307 AAGGCTTTCTTATATGTTTAAAAA 3330
              ||||||||||||||||||||||||
Db       2881 AAGGCTTTCTTATATGTTTAAAAA 2904
```

```
RESULT 18
AAH93867
ID   AAH93867 standard; cDNA; 2904 BP.
XX
AC   AAH93867;
XX
DT   04-OCT-2001  (first entry)
XX
DE   P553S cDNA splice variant P553S-12.
XX
KW   Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW   cytostatic; gene therapy; metastasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200151633-A2.
XX
```

```
PD   19-JUL-2001.
XX
PF   16-JAN-2001; 2001WO-US01574.
XX
PR   14-JAN-2000; 2000US-0483672.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Reed SG;
PI   Kalos MD,  Fanger GR,  Day CH,  Retter MW,  Stolk JA,  Skeiky YAW;
PI   Wang A,  Meagher MJ;
XX
DR   WPI; 2001-425873/45.
XX
PT   New polynucleotide encoding a prostate-specific protein, for
PT   diagnosing, monitoring and treating prostate cancer in a patient and
PT   for use in vaccines -
XX
PS   Claim 1; Page 459-460; 543pp; English.
XX
CC   The present invention describes polynucleotide sequences (I) which encode
CC   prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC   and can be used in vaccine production and gene therapy. (I), (II),
CC   antibodies to (II), fusion proteins comprising (II), and isolated
CC   T cells prepared using (I) or (II) are used treat cancer in a patient.
CC   (I) and the antibodies are also used in the detection of cancer in a
CC   patient. The cancer that is diagnosed or treated is particularly
CC   prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
CC   (I) can be used for monitoring the progression of cancer in a patient.
CC   (I) and (II) can also be used to improve diagnostic and therapeutic
CC   methods for prostate cancer. They can indicate the level of metastasis
CC   as well as the prostate volume. AAH93357 to AAH93944 and AAM01115 to
CC   AAM01318 represent polynucleotide and amino acid sequences used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 2904 BP; 542 A; 875 C; 773 G; 714 T; 0 other;

  Query Match             64.4%;  Score 2196.4;  DB 22;  Length 2904;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches 2560;  Conservative    0;  Mismatches    1;  Indels  343;  Gaps    2;

Qy        770 GTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC 829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC 60

Qy        830 TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTG 889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTG 120

Qy        890 CTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCG 949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCG 180

Qy        950 CTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGT 1009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGT 240

Qy       1010 CCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAG 1069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAG 300

Qy       1070 CTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGG 1129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGG 360

Qy       1130 ATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAG 1189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAG 420

Qy       1190 GGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAA--------- 1240
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGAAGGCCT 480

Qy       1241 ------------------------------------------------------------ 1240

Db        481 CTGGCTGCTCTAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGACAGAAGGAAAGGCGGA 540

Qy       1241 ------------------------------------------------------------ 1240

Db        541 GCTTATTCAAAGTCTAGAGGGAGTGGAGGAGTTAAGGCTGGATTTCAGATCTGCCTGGTT 600

Qy       1241 ------------------------------------------------------------ 1240

Db        601 CCAGCCGCAGTGTGCCCTCTGCTCCCCCAACGACTTTCCAAATAATCTCACCAGCGCCTT 660

Qy       1241 ------------------------------------------------------------ 1240

Db        661 CCAGCTCAGGCGTCCTAGAAGCGTCTTGAAGCCTATGGCCAGCTGTCTTTGTGTTCCCTC 720

Qy       1241 ------------------------------------------------------------ 1240

Db        721 TCACCCGCCTGTCCTCACAGCTGAGACTCCCAGGAAACCTTCAGACTACCTTCCTCTGCC 780

Qy       1241 ---------------------------------GGCGTTCGGATGGCAGCCTGGGGCTGT 1268
                                               |||||||||||||||||||||||||||
Db        781 TTCAGCAAGGGGCGTTGCCCACATTCTCTGAGGGCGTTCGGATGGCAGCCTGGGGCTGT 840

Qy       1269 TCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGAT 1328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGAT 900

Qy       1329 TCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCA 1388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCA 960

Qy       1389 CATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCT 1448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCT 1020

Qy       1449 TCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGG 1508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGG 1080

Qy       1509 TGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGA 1568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGA 1140

Qy       1569 CCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTG 1628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTG 1200

Qy       1629 GAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCT 1688
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCT 1260

Qy       1689 CCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCT 1748
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCT 1320

Qy       1749 GCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGT 1808
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGT 1380

Qy       1809 TTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAG 1868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAG 1440

Qy       1869 GCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGG 1928
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGG 1500

Qy       1929 CCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGT 1988
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGT 1560

Qy       1989 CCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTG 2048
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1561 CCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTG 1620

Qy        2049 CTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACA 2108
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1621 CTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACA 1680

Qy        2109 GCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGG 2168
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1681 GCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGG 1740

Qy        2169 CCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTG 2228
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1741 CCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTG 1800

Qy        2229 GAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGT 2288
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1801 GAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGT 1860

Qy        2289 TGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCA 2348
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1861 TGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCA 1920

Qy        2349 TCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGG 2408
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1921 TCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGG 1980

Qy        2409 ATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTG 2466
               |||||||||||||||||||||||||||||||||  | |||||||||||||||||||||||
Db        1981 ATGAAACACTCCTCCATGGGATTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTG 2040

Qy        2467 AGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCA 2526
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2041 AGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCA 2100

Qy        2527 CCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGA 2586
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2101 CCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGA 2160

Qy        2587 CACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAG 2646
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2161 CACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAG 2220

Qy        2647 CTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCC 2706
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2221 CTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCC 2280

Qy        2707 CTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCC 2766
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2281 CTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCC 2340

Qy        2767 CCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAA 2826
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2341 CCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAA 2400

Qy        2827 TGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAAC 2886
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2401 TGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAAC 2460

Qy        2887 GGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCC 2946
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2461 GGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCC 2520

Qy        2947 TCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAAC 3006
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2521 TCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAAC 2580

Qy        3007 TTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGA 3066
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2581 TTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGA 2640

Qy        3067 CCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCT 3126
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2641 CCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCT 2700

Qy        3127 GTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTC 3186
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2701 GTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTC 2760

Qy        3187 TTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGG 3246
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2761 TTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGG 2820

Qy        3247 GGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTA 3306
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2821 GGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTA 2880

Qy        3307 AAGGCTTTCTTATATGTTTAAAAA 3330
               ||||||||||||||||||||||||
Db        2881 AAGGCTTTCTTATATGTTTAAAAA 2904
```

```
RESULT 19
ABL95410
ID   ABL95410 standard; cDNA; 2904 BP.
XX
AC   ABL95410;
XX
DT   19-JUL-2002  (first entry)
XX
DE   Human P553S splice variant SEQ ID NO 703.
XX
KW   Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
KW   gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   US2002022248-A1.
XX
PD   21-FEB-2002.
XX
PF   12-JAN-2001; 2001US-0759143.
XX
PR   25-FEB-1997;   97US-0806099.
PR   01-AUG-1997;   97US-0904804.
PR   09-FEB-1998;   98US-0020956.
PR   25-FEB-1998;   98US-0030607.
PR   14-JUL-1998;   98US-0115453.
PR   23-SEP-1998;   98US-0159812.
PR   15-JAN-1999;   99US-0232149.
PR   09-APR-1999;   99US-0288946.
PR   13-JUL-1999;   99US-0352616.
PR   12-NOV-1999;   99US-0439313.
PR   18-NOV-1999;   99US-0443686.
PR   14-JAN-2000; 2000US-0483672.
PR   27-MAR-2000; 2000US-0536857.
PR   09-MAY-2000; 2000US-0568100.
PR   12-MAY-2000; 2000US-0570737.
PR   13-JUN-2000; 2000US-0593793.
PR   27-JUN-2000; 2000US-0605783.
PR   10-AUG-2000; 2000US-0636215.
PR   29-AUG-2000; 2000US-0651236.
PR   06-SEP-2000; 2000US-0657279.
PR   02-OCT-2000; 2000US-0679426.
PR   10-OCT-2000; 2000US-0685166.
XX
PA   (XUJJ/) XU J.
PA   (DILL/) DILLON D C.
PA   (MITC/) MITCHAM J L.
PA   (HARL/) HARLOCKER S L.
PA   (JIAN/) JIANG Y.
PA   (KALO/) KALOS M D.
PA   (FANG/) FANGER G R.
PA   (RETT/) RETTER M W.
PA   (STOL/) STOLK J A.
PA   (DAYC/) DAY C H.
```

```
PA   (VEDV/) VEDVICK T S.
PA   (CART/) CARTER D.
PA   (LISX/) LI S X.
PA   (WANG/) WANG A.
PA   (SKEI/) SKEIKY Y A W.
PA   (HEPL/) HEPLER W T.
PA   (HEND/) HENDERSON R A.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Kalos MD;
PI   Fanger GR,  Retter MW,  Stolk JA,  Day CH,  Vedvick TS,  Carter D;
PI   Li SX,  Wang A,  Skeiky YAW,  Hepler WT,  Henderson RA;
XX
DR   WPI; 2002-255649/30.
XX
PT   New prostate-specific polynucleotides for diagnosing and treating
PT   diseases, in particular prostate cancer, and as markers for the
PT   progression of cancer -
XX
PS   Claim 1; SEQ ID NO 703; 87pp; English.
XX
CC   The present invention provides prostate-specific coding sequences and
CC   their encoded proteins. These can be used in the diagnosis and treatment
CC   of cancers, particularly prostate cancer. The present sequence is a cDNA
CC   described in the invention.
XX
SQ   Sequence 2904 BP; 542 A; 875 C; 773 G; 714 T; 0 other;

  Query Match             64.4%;  Score 2196.4;  DB 24;  Length 2904;
  Best Local Similarity   88.2%;  Pred. No. 0;
  Matches 2560;  Conservative    0;  Mismatches    1;  Indels  343;  Gaps    2;

Qy        770 GTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC 829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC 60

Qy        830 TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTG 889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTG 120

Qy        890 CTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCG 949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCG 180

Qy        950 CTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGT 1009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGT 240

Qy       1010 CCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAG 1069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAG 300

Qy       1070 CTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGG 1129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGG 360

Qy       1130 ATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAG 1189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAG 420

Qy       1190 GGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAA--------- 1240
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGAAGGCCT 480

Qy       1241 ------------------------------------------------------------ 1240

Db        481 CTGGCTGCTCTAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGACAGAAGGAAAGGCGGA 540

Qy       1241 ------------------------------------------------------------ 1240

Db        541 GCTTATTCAAAGTCTAGAGGGAGTGGAGGAGTTAAGGCTGGATTTCAGATCTGCCTGGTT 600

Qy       1241 ------------------------------------------------------------ 1240

Db        601 CCAGCCGCAGTGTGCCCTCTGCTCCCCCAACGACTTTCCAAATAATCTCACCAGCGCCTT 660

Qy       1241 ------------------------------------------------------------ 1240

Db        661 CCAGCTCAGGCGTCCTAGAAGCGTCTTGAAGCCTATGGCCAGCTGTCTTTGTGTTCCCTC 720

Qy       1241 ------------------------------------------------------------ 1240

Db        721 TCACCCGCCTGTCCTCACAGCTGAGACTCCCAGGAAACCTTCAGACTACCTTCCTCTGCC 780

Qy       1241 --------------------------------GGCGTTCGGATGGGCAGCCTGGGGCTGT 1268
                                              ||||||||||||||||||||||||||||
Db        781 TTCAGCAAGGGGCGTTGCCCACATTCTCTGAGGGCGTTCGGATGGGCAGCCTGGGGCTGT 840

Qy       1269 TCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGAT 1328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGAT 900

Qy       1329 TCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCA 1388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCA 960

Qy       1389 CATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCT 1448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCT 1020

Qy       1449 TCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGG 1508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGG 1080

Qy       1509 TGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGA 1568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGA 1140

Qy       1569 CCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTG 1628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTG 1200

Qy       1629 GAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCT 1688
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCT 1260

Qy       1689 CCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCT 1748
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCT 1320

Qy       1749 GCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGT 1808
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGT 1380

Qy       1809 TTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAG 1868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAG 1440

Qy       1869 GCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGG 1928
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGG 1500

Qy       1929 CCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGT 1988
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGT 1560

Qy       1989 CCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTG 2048
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTG 1620

Qy       2049 CTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACA 2108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACA 1680

Qy       2109 GCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGG 2168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGG 1740
```

```
Qy       2169 CCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTG 2228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTG 1800

Qy       2229 GAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGT 2288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGT 1860

Qy       2289 TGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCA 2348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCA 1920

Qy       2349 TCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGG 2408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGG 1980

Qy       2409 ATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTG 2466
              |||||||||||||||||||||||||||||||||     |   ||||||||||||||||||
Db       1981 ATGAAACACTCCTCCATGGGATTTGAACATATGAAGTTATTTGTAGGGGAAGAGTCCTG 2040

Qy       2467 AGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCA 2526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 AGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCA 2100

Qy       2527 CCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGA 2586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGA 2160

Qy       2587 CACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAG 2646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 CACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAG 2220

Qy       2647 CTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCC 2706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCC 2280

Qy       2707 CTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCC 2766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCC 2340

Qy       2767 CCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAA 2826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 CCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAA 2400

Qy       2827 TGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAAC 2886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAAC 2460

Qy       2887 GGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCC 2946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCC 2520

Qy       2947 TCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAAC 3006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAAC 2580

Qy       3007 TTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGA 3066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 TTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGA 2640

Qy       3067 CCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCT 3126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCT 2700

Qy       3127 GTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTC 3186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTC 2760

Qy       3187 TTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGG 3246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGG 2820

Qy       3247 GGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTA 3306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 GGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTA 2880

Qy       3307 AAGGCTTTCTTATATGTTTAAAAA 3330
              ||||||||||||||||||||||||
Db       2881 AAGGCTTTCTTATATGTTTAAAAA 2904
```

```
RESULT 20
AAS64038
ID   AAS64038 standard; cDNA; 4894 BP.
XX
AC   AAS64038;
XX
DT   29-JAN-2002  (first entry)
XX
DE   Human prosate cDNA P553S splice variant #1.
XX
KW   Human; prostate cancer; ss; cytostatic; immunostimulant; tumour.
XX
OS   Homo sapiens.
XX
PN   WO200173032-A2.
XX
PD   04-OCT-2001.
XX
PF   27-MAR-2001; 2001WO-US09919.
XX
PR   27-MAR-2000; 2000US-0536857.
PR   09-MAY-2000; 2000US-0568100.
PR   12-MAY-2000; 2000US-0570737.
PR   13-JUN-2000; 2000US-0593793.
PR   27-JUN-2000; 2000US-0605783.
PR   10-AUG-2000; 2000US-0636215.
PR   29-AUG-2000; 2000US-0651236.
PR   06-SEP-2000; 2000US-0657279.
PR   02-OCT-2000; 2000US-0679426.
PR   10-OCT-2000; 2000US-0685166.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Kalos MD;
PI   Fanger GR,  Retter MW,  Stolk JA,  Day CH,  Vedvick TS,  Carter D;
PI   Li SX,  Wang A,  Skeiky YAW,  Hepler WT,  Henderson RA;
XX
DR   WPI; 2001-639232/73.
DR   N-PSDB; AAU69874, AAU69875.
XX
PT   New human prostate-specific polypeptides and polynucleotides useful for
PT   the diagnosis and treatment of cancer, especially prostate cancer -
XX
PS   Claim 1; Page 480-482; 579pp; English.
XX
CC   The invention relates to isolated prostate-specific
CC   polynucleotides, polypeptides, fusion proteins of the polypeptides,
CC   antibodies raised against the polypeptides (or antigenic epitopes
CC   derived from them) and antigen-presenting cells expressing the
CC   polypeptides. The antibodies are useful for detecting the presence of
CC   cancer, especially prostate cancer. The polypeptides, polynucleotides and
CC   the antigen-presenting cells are useful for stimulating and/or expanding
CC   T cells specific for a tumour protein, and for inhibiting the development
CC   of cancer especially prostate cancer. Compositions comprising the
CC   polynucleotide and/or polypeptide are useful for stimulating an immune
CC   response, and for treating cancer. The oligonucleotide is useful for
CC   detecting cancer. The present sequence is a prostate specific
CC   polynucleotide of the invention.
XX
SQ   Sequence 4894 BP; 928 A; 1448 C; 1354 G; 1163 T; 1 other;

  Query Match             62.8%;  Score 2142.8;  DB 22;  Length 4894;
  Best Local Similarity   80.1%;  Pred. No. 0;
  Matches 2872;  Conservative    1;  Mismatches    3;  Indels  708;  Gaps    2;
```

```
Qy    455 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 514
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1311 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 1370

Qy    515 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 574
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1371 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 1430

Qy    575 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 634
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1431 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 1490

Qy    635 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 694
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1491 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 1550

Qy    695 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 754
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1551 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 1610

Qy    755 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 814
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1611 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 1670

Qy    815 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 874
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1671 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 1730

Qy    875 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 934
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1731 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 1790

Qy    935 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 994
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1791 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 1850

Qy    995 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1054
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1851 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1910

Qy   1055 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1114
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1911 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1970

Qy   1115 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 1174
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1971 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 2030

Qy   1175 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 1234
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2031 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 2090

Qy   1235 GATGA------------------------------------------------------- 1239
          |||||
Db   2091 GATGAAGGTAAGGCCTTGGCAGCCAGCAGAGGCTGGTGTGGAGCCGCCCACCAGAGACG 2150

Qy   1240 ------------------------------------------------------------ 1239

Db   2151 ACACTCGGGGCTGTGTCTGGGCTGGTGCCTCTCCATCCTGGCCCCGACTTCTCTGTCAGG 2210

Qy   1240 ------------------------------------------------------------ 1239

Db   2211 AAAGTGGGGATGGACCCCATCTGCATACACGGCTTCTCATGGGTGTGGAACATCTCTGCT 2270

Qy   1240 ------------------------------------------------------------ 1239

Db   2271 TGCGGTTTCAGGAAGGCCTCTGGCTGCTCTAGGAGTCTGATCAGAGTCGTTGCCCCAGTT 2330

Qy   1240 ------------------------------------------------------------ 1239

Db   2331 TGACAGAAGGAAAGGCGGAGCTTATTCAAAGTCTAGAGGGAGTGGAGGAGTTAAGGCTGG 2390

Qy   1240 ------------------------------------------------------------ 1239

Db   2391 ATTTCAGATCTGCCTGGTTCCAGCCGCAGTGTGCCCTCTGCTCCCCCAACGACTTTCCAA 2450

Qy   1240 ------------------------------------------------------------ 1239

Db   2451 ATAATCTCACCAGCGCCTTCCAGCTCAGGCGTCCTAGAAGCGTCTTGAAGCCTATGGCCA 2510

Qy   1240 ------------------------------------------------------------ 1239

Db   2511 GCTGTCTTTGTGTTCCCTCTCACCCGCCTGTCCTCACAGCTGAGACTCCCAGGAAACCTT 2570

Qy   1240 ------------------------------------------------------------ 1239

Db   2571 CAGACTACCTTCCTCTGCCTTCAGCAAGGGGCGTTGCCCACATTCTCTGAGGGTCAGTGG 2630

Qy   1240 ------------------------------------------------------------ 1239

Db   2631 AAGAACCTAGACTCCCATTGCTAGAGGTAGAAAGGGGAAGGGTGCTGGGGAGCAGGGCTG 2690

Qy   1240 ------------------------------------------------------------ 1239

Db   2691 GTCCACAGCAGGTCTCGTGCAGCAGGTACCTGTGGTTCCGCCTTCTCATCTCCCTGAGAC 2750

Qy   1240 ---------------------------------------------------AGGCGTTCG 1248
                                                             |||||||||
Db   2751 TGCTCCGACCCTTCCCTCCCAGGCTCTGTCTGATGGCCCCTCTCCCTCTGCAGGCGTTCG 2810

Qy   1249 GATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCAT 1308
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2811 GATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCAT 2870

Qy   1309 GGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 1368
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2871 GGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 2930

Qy   1369 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 1428
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2931 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 2990

Qy   1429 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 1488
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2991 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 3050

Qy   1489 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 1548
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3051 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 3110

Qy   1549 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 1608
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3111 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 3170

Qy   1609 TAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 1668
          |||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db   3171 TAATGGACACGTGGGTGYTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 3230

Qy   1669 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 1728
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3231 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 3290

Qy   1729 GGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTC 1788
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3291 GGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTC 3350

Qy   1789 CCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGC 1848
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3351 CCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGC 3410

Qy   1849 CTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT 1908
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3411 CTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT 3470

Qy   1909 ATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG 1968
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3471 ATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG 3530

Qy         1969 GAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGG 2028
                |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db         3531 GAGGGCCTGCCTCACTGGGTCCCAGCTCCCTGCTCCTGTTAGCCCCATGGGGCTGCCGGG 3590

Qy         2029 CTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTG 2088
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3591 CTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTG 3650

Qy         2089 CTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTC 2148
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3651 CTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTC 3710

Qy         2149 TAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGC 2208
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3711 TAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGC 3770

Qy         2209 CAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG 2268
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3771 CAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG 3830

Qy         2269 TTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAA 2328
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3831 TTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAA 3890

Qy         2329 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCT 2388
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3891 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCT 3950

Qy         2389 CTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTA 2446
                |||||||||||||||||||||||||||||||| ||||||||||||||||||||   | |||
Db         3951 CTTGCATGGGAGTTTCTAGGATGAAACACTCCACCATGGGATTTGAACATATGAAAGTTA 4010

Qy         2447 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 2506
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4011 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 4070

Qy         2507 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 2566
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4071 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 4130

Qy         2567 TTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGT 2626
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4131 TTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGT 4190

Qy         2627 AGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 2686
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4191 AGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 4250

Qy         2687 ATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCT 2746
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4251 ATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCT 4310

Qy         2747 TCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCA 2806
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4311 TCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCA 4370

Qy         2807 TCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGG 2866
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4371 TCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGG 4430

Qy         2867 GTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGT 2926
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4431 GTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGT 4490

Qy         2927 TCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 2986
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4491 TCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 4550

Qy         2987 AAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCC 3046
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4551 AAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCC 4610

Qy         3047 TGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATC 3106
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4611 TGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATC 4670

Qy         3107 TCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCC 3166
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4671 TCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCC 4730

Qy         3167 TGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATG 3226
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4731 TGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATG 4790

Qy         3227 TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAA 3286
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4791 TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAA 4850

Qy         3287 TGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAA 3330
                ||||||||||||||||||||||||||||||||||||||||||||
Db         4851 TGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAA 4894
```

```
RESULT 21
AAH93866
ID   AAH93866 standard; cDNA; 4894 BP.
XX
AC   AAH93866;
XX
DT   04-OCT-2001  (first entry)
XX
DE   P553S cDNA splice variant P553S-14.
XX
KW   Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW   cytostatic; gene therapy; metastasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200151633-A2.
XX
PD   19-JUL-2001.
XX
PF   16-JAN-2001; 2001WO-US01574.
XX
PR   14-JAN-2000; 2000US-0483672.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Reed SG;
PI   Kalos MD,  Fanger GR,  Day CH,  Retter MW,  Stolk JA,  Skeiky YAW;
PI   Wang A,  Meagher MJ;
XX
DR   WPI; 2001-425873/45.
XX
PT   New polynucleotide encoding a prostate-specific protein, for
PT   diagnosing, monitoring and treating prostate cancer in a patient and
PT   for use in vaccines -
XX
PS   Claim 1; Page 457-459; 543pp; English.
XX
CC   The present invention describes polynucleotide sequences (I) which encode
CC   prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC   and can be used in vaccine production and gene therapy. (I), (II),
CC   antibodies to (II), fusion proteins comprising (II), and isolated
CC   T cells prepared using (I) or (II) are used treat cancer in a patient.
CC   (I) and the antibodies are also used in the detection of cancer in a
CC   patient. The cancer that is diagnosed or treated is particularly
CC   prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
CC   (I) can be used for monitoring the progression of cancer in a patient.
CC   (I) and (II) can also be used to improve diagnostic and therapeutic
CC   methods for prostate cancer. They can indicate the level of metastasis
CC   as well as the prostate volume. AAH93357 to AAH93944 and AAM01115 to
CC   AAM01318 represent polynucleotide and amino acid sequences used in the
```

```
CC   exemplification of the present invention.
XX
SQ   Sequence 4894 BP; 928 A; 1448 C; 1354 G; 1163 T; 1 other;

  Query Match             62.8%;  Score 2142.8;  DB 22;  Length 4894;
  Best Local Similarity   80.1%;  Pred. No. 0;
  Matches 2872;  Conservative    1;  Mismatches    3;  Indels  708;  Gaps    2;

Qy        455 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1311 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 1370

Qy        515 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1371 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 1430

Qy        575 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1431 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 1490

Qy        635 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 694
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1491 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 1550

Qy        695 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1551 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 1610

Qy        755 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 814
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1611 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 1670

Qy        815 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 874
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1671 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 1730

Qy        875 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 934
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1731 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 1790

Qy        935 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 994
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1791 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 1850

Qy        995 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1054
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1851 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1910

Qy       1055 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1911 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1970

Qy       1115 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 1174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1971 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 2030

Qy       1175 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 1234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2031 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 2090

Qy       1235 GATGA------------------------------------------------------- 1239
              |||||
Db       2091 GATGAAGGTAAGGCCTTGGCAGCCAGCAGAGGCTGGTGTGGGAGCCGCCCACCAGAGACG 2150

Qy       1240 ------------------------------------------------------------ 1239

Db       2151 ACACTCGGGGCTGTGTCTGGGCTGGTGCCTCTCCATCCTGGCCCCGACTTCTCTGTCAGG 2210

Qy       1240 ------------------------------------------------------------ 1239

Db       2211 AAAGTGGGGATGGACCCCATCTGCATACACGGCTTCTCATGGGTGTGGAACATCTCTGCT 2270

Qy       1240 ------------------------------------------------------------ 1239

Db       2271 TGCGGTTTCAGGAAGGCCTCTGGCTGCTCTAGGAGTCTGATCAGAGTCGTTGCCCCAGTT 2330

Qy       1240 ------------------------------------------------------------ 1239

Db       2331 TGACAGAAGGAAAGGCGGAGCTTATTCAAAGTCTAGAGGGAGTGGAGGAGTTAAGGCTGG 2390

Qy       1240 ------------------------------------------------------------ 1239

Db       2391 ATTTCAGATCTGCCTGGTTCCAGCCGCAGTGTGCCCTCTGCTCCCCCAACGACTTTCCAA 2450

Qy       1240 ------------------------------------------------------------ 1239

Db       2451 ATAATCTCACCAGCGCCTTCCAGCTCAGGCGTCCTAGAAGCGTCTTGAAGCCTATGGCCA 2510

Qy       1240 ------------------------------------------------------------ 1239

Db       2511 GCTGTCTTTGTGTTCCCTCTCACCCGCCTGTCCTCACAGCTGAGACTCCCAGGAAACCTT 2570

Qy       1240 ------------------------------------------------------------ 1239

Db       2571 CAGACTACCTTCCTCTGCCTTCAGCAAGGGGCGTTGCCCACATTCTCTGAGGGTCAGTGG 2630

Qy       1240 ------------------------------------------------------------ 1239

Db       2631 AAGAACCTAGACTCCCATTGCTAGAGGTAGAAAGGGGAAGGGTGCTGGGGAGCAGGGCTG 2690

Qy       1240 ------------------------------------------------------------ 1239

Db       2691 GTCCACAGCAGGTCTCGTGCAGCAGGTACCTGTGGTTCCGCCTTCTCATCTCCCTGAGAC 2750

Qy       1240 ---------------------------------------------------AGGCGTTCG 1248
                                                                 |||||||||
Db       2751 TGCTCCGACCCTTCCCTCCCAGGCTCTGTCTGATGGCCCCTCTCCCTCTGCAGGCGTTCG 2810

Qy       1249 GATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCAT 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2811 GATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCAT 2870

Qy       1309 GGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 1368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2871 GGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 2930

Qy       1369 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 1428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2931 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 2990

Qy       1429 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 1488
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2991 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 3050

Qy       1489 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 1548
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3051 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 3110

Qy       1549 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 1608
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3111 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 3170

Qy       1609 TAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 1668
              |||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db       3171 TAATGGACACGTGGGTGYTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 3230

Qy       1669 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 1728
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3231 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 3290

Qy       1729 GGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTC 1788
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3291 GGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTC 3350

Qy       1789 CCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGC 1848
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        3351 CCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGC 3410

Qy        1849 CTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT 1908
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3411 CTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT 3470

Qy        1909 ATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG 1968
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3471 ATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG 3530

Qy        1969 GAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGG 2028
               |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db        3531 GAGGGCCTGCCTCACTGGGTCCCAGCTCCCTGCTCCTGTTAGCCCCATGGGGCTGCCGGG 3590

Qy        2029 CTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTG 2088
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3591 CTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTG 3650

Qy        2089 CTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTC 2148
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3651 CTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTC 3710

Qy        2149 TAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGC 2208
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3711 TAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGC 3770

Qy        2209 CAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG 2268
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3771 CAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG 3830

Qy        2269 TTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAA 2328
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3831 TTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAA 3890

Qy        2329 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCT 2388
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3891 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCT 3950

Qy        2389 CTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTA 2446
               |||||||||||||||||||||||||||||||| ||||||||||||||||||||   | |||
Db        3951 CTTGCATGGGAGTTTCTAGGATGAAACACTCCACCATGGGATTTGAACATATGAAGTTA 4010

Qy        2447 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 2506
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4011 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 4070

Qy        2507 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 2566
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4071 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 4130

Qy        2567 TTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGT 2626
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4131 TTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGT 4190

Qy        2627 AGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 2686
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4191 AGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 4250

Qy        2687 ATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCT 2746
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4251 ATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCT 4310

Qy        2747 TCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCA 2806
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4311 TCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCA 4370

Qy        2807 TCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGG 2866
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4371 TCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGG 4430

Qy        2867 GTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGT 2926
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4431 GTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGT 4490

Qy        2927 TCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 2986
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4491 TCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 4550

Qy        2987 AAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCC 3046
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4551 AAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCC 4610

Qy        3047 TGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATC 3106
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4611 TGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATC 4670

Qy        3107 TCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCC 3166
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4671 TCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCC 4730

Qy        3167 TGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATG 3226
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4731 TGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATG 4790

Qy        3227 TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAA 3286
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4791 TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAA 4850

Qy        3287 TGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAA 3330
               ||||||||||||||||||||||||||||||||||||||||||||
Db        4851 TGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAA 4894
```

```
RESULT 22
ABL95409
ID   ABL95409 standard; cDNA; 4894 BP.
XX
AC   ABL95409;
XX
DT   19-JUL-2002  (first entry)
XX
DE   Human P553S splice variant SEQ ID NO 702.
XX
KW   Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
KW   gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   US2002022248-A1.
XX
PD   21-FEB-2002.
XX
PF   12-JAN-2001; 2001US-0759143.
XX
PR   25-FEB-1997;   97US-0806099.
PR   01-AUG-1997;   97US-0904804.
PR   09-FEB-1998;   98US-0020956.
PR   25-FEB-1998;   98US-0030607.
PR   14-JUL-1998;   98US-0115453.
PR   23-SEP-1998;   98US-0159812.
PR   15-JAN-1999;   99US-0232149.
PR   09-APR-1999;   99US-0288946.
PR   13-JUL-1999;   99US-0352616.
PR   12-NOV-1999;   99US-0439313.
PR   18-NOV-1999;   99US-0443686.
PR   14-JAN-2000; 2000US-0483672.
PR   27-MAR-2000; 2000US-0536857.
PR   09-MAY-2000; 2000US-0568100.
PR   12-MAY-2000; 2000US-0570737.
PR   13-JUN-2000; 2000US-0593793.
PR   27-JUN-2000; 2000US-0605783.
PR   10-AUG-2000; 2000US-0636215.
PR   29-AUG-2000; 2000US-0651236.
PR   06-SEP-2000; 2000US-0657279.
PR   02-OCT-2000; 2000US-0679426.
PR   10-OCT-2000; 2000US-0685166.
```

```
XX
PA   (XUJJ/) XU J.
PA   (DILL/) DILLON D C.
PA   (MITC/) MITCHAM J L.
PA   (HARL/) HARLOCKER S L.
PA   (JIAN/) JIANG Y.
PA   (KALO/) KALOS M D.
PA   (FANG/) FANGER G R.
PA   (RETT/) RETTER M W.
PA   (STOL/) STOLK J A.
PA   (DAYC/) DAY C H.
PA   (VEDV/) VEDVICK T S.
PA   (CART/) CARTER D.
PA   (LISX/) LI S X.
PA   (WANG/) WANG A.
PA   (SKEI/) SKEIKY Y A W.
PA   (HEPL/) HEPLER W T.
PA   (HEND/) HENDERSON R A.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Kalos MD;
PI   Fanger GR,  Retter MW,  Stolk JA,  Day CH,  Vedvick TS,  Carter D;
PI   Li SX,  Wang A,  Skeiky YAW,  Hepler WT,  Henderson RA;
XX
DR   WPI; 2002-255649/30.
XX
PT   New prostate-specific polynucleotides for diagnosing and treating
PT   diseases, in particular prostate cancer, and as markers for the
PT   progression of cancer  -
XX
PS   Claim 1; SEQ ID NO 702; 87pp; English.
XX
CC   The present invention provides prostate-specific coding sequences and
CC   their encoded proteins. These can be used in the diagnosis and treatment
CC   of cancers, particularly prostate cancer. The present sequence is a cDNA
CC   described in the invention.
XX
SQ   Sequence 4894 BP; 928 A; 1448 C; 1354 G; 1163 T; 1 other;

  Query Match             62.8%;  Score 2142.8;  DB 24;  Length 4894;
  Best Local Similarity   80.1%;  Pred. No. 0;
  Matches 2872;  Conservative    1;  Mismatches    3;  Indels  708;  Gaps    2;

Qy        455 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1311 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 1370

Qy        515 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1371 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 1430

Qy        575 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1431 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 1490

Qy        635 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 694
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1491 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 1550

Qy        695 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1551 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 1610

Qy        755 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 814
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1611 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 1670

Qy        815 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 874
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1671 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 1730

Qy        875 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 934
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1731 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 1790

Qy        935 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 994
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1791 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 1850

Qy        995 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1054
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1851 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1910

Qy       1055 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1911 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1970

Qy       1115 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 1174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1971 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 2030

Qy       1175 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 1234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2031 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 2090

Qy       1235 GATGA------------------------------------------------------- 1239
              |||||
Db       2091 GATGAAGGTAAGGCCTTGGCAGCCAGCAGAGGCTGGTGTGGGAGCCGCCCACCAGAGACG 2150

Qy       1240 ------------------------------------------------------------ 1239

Db       2151 ACACTCGGGGCTGTGTCTGGGCTGGTGCCTCTCCATCCTGGCCCCGACTTCTCTGTCAGG 2210

Qy       1240 ------------------------------------------------------------ 1239

Db       2211 AAAGTGGGGATGGACCCCATCTGCATACACGGCTTCTCATGGGTGTGGAACATCTCTGCT 2270

Qy       1240 ------------------------------------------------------------ 1239

Db       2271 TGCGGTTTCAGGAAGGCCTCTGGCTGCTCTAGGAGTCTGATCAGAGTCGTTGCCCCAGTT 2330

Qy       1240 ------------------------------------------------------------ 1239

Db       2331 TGACAGAAGGAAAGGCGGAGCTTATTCAAAGTCTAGAGGGAGTGGAGGAGTTAAGGCTGG 2390

Qy       1240 ------------------------------------------------------------ 1239

Db       2391 ATTTCAGATCTGCCTGGTTCCAGCCGCAGTGTGCCCTCTGCTCCCCCAACGACTTTCCAA 2450

Qy       1240 ------------------------------------------------------------ 1239

Db       2451 ATAATCTCACCAGCGCCTTCCAGCTCAGGCGTCCTAGAAGCGTCTTGAAGCCTATGGCCA 2510

Qy       1240 ------------------------------------------------------------ 1239

Db       2511 GCTGTCTTTGTGTTCCCTCTCACCCGCCTGTCCTCACAGCTGAGACTCCCAGGAAACCTT 2570

Qy       1240 ------------------------------------------------------------ 1239

Db       2571 CAGACTACCTTCCTCTGCCTTCAGCAAGGGGCGTTGCCCACATTCTCTGAGGGTCAGTGG 2630

Qy       1240 ------------------------------------------------------------ 1239

Db       2631 AAGAACCTAGACTCCCATTGCTAGAGGTAGAAAGGGGAAGGGTGCTGGGGAGCAGGGCTG 2690

Qy       1240 ------------------------------------------------------------ 1239

Db       2691 GTCCACAGCAGGTCTCGTGCAGCAGGTACCTGTGGTTCCGCCTTCTCATCTCCCTGAGAC 2750

Qy       1240 ---------------------------------------------------AGGCGTTCG 1248
                                                                 |||||||||
Db       2751 TGCTCCGACCCTTCCCTCCCAGGCTCTGTCTGATGGCCCCTCTCCCTCTGCAGGCGTTCG 2810

Qy       1249 GATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCAT 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2811 GATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCAT 2870
```

```
Qy    1309 GGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 1368
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2871 GGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 2930

Qy    1369 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 1428
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2931 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 2990

Qy    1429 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 1488
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2991 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 3050

Qy    1489 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 1548
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3051 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 3110

Qy    1549 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 1608
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3111 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 3170

Qy    1609 TAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 1668
           |||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db    3171 TAATGGACACGTGGGTGYTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 3230

Qy    1669 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 1728
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3231 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 3290

Qy    1729 GGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTC 1788
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3291 GGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTC 3350

Qy    1789 CCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGC 1848
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3351 CCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGC 3410

Qy    1849 CTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT 1908
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3411 CTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT 3470

Qy    1909 ATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG 1968
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3471 ATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG 3530

Qy    1969 GAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGG 2028
           |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db    3531 GAGGGCCTGCCTCACTGGGTCCCAGCTCCCTGCTCCTGTTAGCCCCATGGGGCTGCCGGG 3590

Qy    2029 CTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTG 2088
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3591 CTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTG 3650

Qy    2089 CTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTC 2148
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3651 CTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTC 3710

Qy    2149 TAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGC 2208
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3711 TAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGC 3770

Qy    2209 CAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG 2268
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3771 CAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG 3830

Qy    2269 TTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAA 2328
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3831 TTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAA 3890

Qy    2329 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCT 2388
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3891 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCT 3950

Qy    2389 CTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTA 2446
           |||||||||||||||||||||||||||||||| |||||||||||||||||||||  | ||
Db    3951 CTTGCATGGGAGTTTCTAGGATGAAACACTCCACCATGGGATTTGAACATATGAAAGTTA 4010

Qy    2447 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 2506
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4011 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 4070

Qy    2507 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 2566
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4071 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 4130

Qy    2567 TTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGT 2626
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4131 TTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGT 4190

Qy    2627 AGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 2686
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4191 AGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 4250

Qy    2687 ATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCT 2746
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4251 ATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCT 4310

Qy    2747 TCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCA 2806
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4311 TCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCA 4370

Qy    2807 TCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGG 2866
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4371 TCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGG 4430

Qy    2867 GTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGT 2926
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4431 GTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGT 4490

Qy    2927 TCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 2986
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4491 TCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 4550

Qy    2987 AAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCC 3046
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4551 AAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCC 4610

Qy    3047 TGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATC 3106
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4611 TGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATC 4670

Qy    3107 TCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCC 3166
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4671 TCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCC 4730

Qy    3167 TGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATG 3226
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4731 TGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATG 4790

Qy    3227 TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAA 3286
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4791 TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAA 4850

Qy    3287 TGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAA 3330
           ||||||||||||||||||||||||||||||||||||||||||||
Db    4851 TGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAA 4894
```

```
RESULT 23
AAD05230
ID   AAD05230 standard; cDNA; 3878 BP.
XX
AC   AAD05230;
XX
DT   18-JUL-2001  (first entry)
XX
DE   Human secreted protein-encoding gene 11 cDNA clone HWBAR14, SEQ ID NO:21.
```

```
XX
KW   Human; secreted protein; proliferative disorder; cancer; tumour; asthma;
KW   foetal abnormality; developmental abnormality; haematopoietic disorder;
KW   immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW   Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;
KW   psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;
KW   inflammation; neurological disorder; Alzheimer's disease; food additive;
KW   angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;
KW   pregnancy-related disorder; endocrine disorder; infection; wound healing;
KW   cell culture; chemotaxis; vulnerary; binding partner identification;
KW   gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             152..1267
FT                   /*tag= a
FT                   /product= "Human secreted protein"
FT                   /transl_except= (pos:209..211, aa:Xaa)
FT                   /note= "Xaa corresponds to any of the naturally occurring
FT                   L-amino acids"
FT   sig_peptide     152..295
FT                   /*tag= b
FT   mat_peptide     296..1264
FT                   /*tag= c
FT                   /product= "Mature human secreted protein"
XX
PN   WO200134629-A1.
XX
PD   17-MAY-2001.
XX
PF   08-NOV-2000; 2000WO-US30654.
XX
PR   12-NOV-1999;   99US-0164835.
PR   27-JUL-2000; 2000US-0221142.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Komatsoulis GA,  Wei P,  Fiscella M,  Baker KP;
XX
DR   WPI; 2001-308779/32.
DR   P-PSDB; AAE01362.
XX
PT   New nucleic acid encoding one of 21 human secreted proteins for
PT   diagnosing, preventing, treating or ameliorating medical conditions,
PT   such as autoimmune disease and cancer, and used as a food additive or
PT   preservative -
XX
PS   Claim 1; Page 388-389; 490pp; English.
XX
CC   AAD05220-AAD05282 represent cDNAs corresponding to 21 human secreted
CC   protein genes, and AAE01352-AAE01413 represent the proteins they encode.
CC   AAE01415-AAE01433 represent human secreted protein fragments or variants.
CC   The secreted proteins and their genes are useful for preventing, treating
CC   or ameliorating medical conditions, e.g., by protein or gene therapy.
CC   Pathological conditions can be diagnosed by determining the amount of the
CC   new protein in a sample or by determining the presence of mutations in
CC   the new genes. Specific uses are described for each of the 21 genes,
CC   based on the tissues in which they are most highly expressed, and include
CC   developing products for the diagnosis or treatment of proliferative
CC   disorders, cancer, tumours, foetal and developmental abnormalities,
CC   haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
CC   diseases (e.g., rheumatoid arthritis), inflammation, allergies,
CC   neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
CC   cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
CC   psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
CC   angiogenic disorders, kidney disorders, gastrointestinal disorders,
CC   pregnancy-related disorders, endocrine disorders, and infections. The
CC   proteins can also be used to aid wound healing and epithelial cell
CC   proliferation, to prevent skin aging due to sunburn, to maintain organs
CC   before transplantation, for supporting cell culture of primary tissues,
CC   to regenerate tissues, to identify their cognate ligands or binding
CC   partners, and in chemotaxis, and can be used as a food additive or
CC   preservative to modify storage properties. Antibodies specific for a
CC   protein of the invention can be used in alleviating symptoms associated
CC   with the disorders mentioned above, and in diagnostic immunoassays e.g.,
CC   radioimmunoassay or enzyme linked immunosorbent assay (ELISA).
CC   The present sequence represents a human secreted protein-encoding cDNA of
CC   the invention.
XX
SQ   Sequence 3878 BP; 709 A; 1164 C; 1044 G; 951 T; 10 other;

  Query Match             62.7%;  Score 2136.8;  DB 22;  Length 3878;
  Best Local Similarity   80.0%;  Pred. No. 0;
  Matches 2885;  Conservative    5;  Mismatches    4;  Indels  710;  Gaps    4;

Qy        455 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 514
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db        245 GGCATTGGTCCAGTGCTGGGCCTGGTNTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 304

Qy        515 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 574
              |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 TGGCGWGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 364

Qy        575 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 424

Qy        635 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 694
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 484

Qy        695 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        485 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 544

Qy        755 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 814
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        545 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 604

Qy        815 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 874
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        605 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 664

Qy        875 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 934
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        665 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 724

Qy        935 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 994
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        725 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 784

Qy        995 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1054
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        785 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 844

Qy       1055 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        845 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 904

Qy       1115 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 1174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        905 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 964

Qy       1175 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 1234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        965 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 1024

Qy       1235 GATGA------------------------------------------------------- 1239
              |||||
Db       1025 GATGAAGGTAAGGCCTTGGCAGCCAGCAGAGGCTGGTGTGGGAGCCGCCCACCAGAGACG 1084

Qy       1240 ------------------------------------------------------------ 1239

Db       1085 ACACTCGGGGCTGTGTCTGGGCTGGTGCCTCTCCATCCTGGCCCCGACTTCTCTGTCAGG 1144
```

```
Qy      1240 ------------------------------------------------------------ 1239

Db      1145 AAAGTGGGGATGGACCCCATCTGCATACACGGCTTCTCATGGGTGTGGAACATCTCTGCT 1204

Qy      1240 ------------------------------------------------------------ 1239

Db      1205 TGCGGTTTCAGGAAGGCCTCTGGCTGCTCTAGGAGTCTGATCAGAGTCGTTGCCCCAGTT 1264

Qy      1240 ------------------------------------------------------------ 1239

Db      1265 TGACAGAAGGAAAGGCGGAGCTTATTCAAAGTCTAGAGGGAGTGGAGGAGTTAAGGCTGG 1324

Qy      1240 ------------------------------------------------------------ 1239

Db      1325 ATTTCAGATCTGCCTGGTTCCAGCCGCAGTGTGCCCTCTGCTCCCCCAACGACTTTCCAA 1384

Qy      1240 ------------------------------------------------------------ 1239

Db      1385 ATAATCTCACCAGCGCCTTCCAGCTCAGGCGTCCTAGAAGCGTCTTGAAGCCTATGGCCA 1444

Qy      1240 ------------------------------------------------------------ 1239

Db      1445 GCTGTCTTTGTGTTCCCTCTCACCCGCCTGTCCTCACAGCTGAGACTCCCAGGAAACCTT 1504

Qy      1240 ------------------------------------------------------------ 1239

Db      1505 CAGACTACCTTCCTCTGCCTTCAGCAAGGGGCGTTGCCCACATTCTCTGAGGGTCAGTGG 1564

Qy      1240 ------------------------------------------------------------ 1239

Db      1565 AAGAACCTAGACTCCCATTGCTAGAGGTAGAAAGGGGAAGGGTGCTGGGGAGCAGGGCTG 1624

Qy      1240 ------------------------------------------------------------ 1239

Db      1625 GTCCACAGCAGGTCTCGTGCAGCAGGTACCTGTGGTTCCGCCTTCTCATCTCCCTGAGAC 1684

Qy      1240 ---------------------------------------------------AGGCGTTCG 1248
                                                                |||||||||
Db      1685 TGCTCCGACCCTTCCCTCCCAGGCTCTGTCTGATGGCCCCTCTCCCTCTGCAGGCGTTCG 1744

Qy      1249 GATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCAT 1308
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1745 GATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCAT 1804

Qy      1309 GGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 1368
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1805 GGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 1864

Qy      1369 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 1428
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1865 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 1924

Qy      1429 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 1488
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1925 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 1984

Qy      1489 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 1548
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1985 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 2044

Qy      1549 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 1608
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2045 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 2104

Qy      1609 TAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 1668
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2105 TAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 2164

Qy      1669 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 1728
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2165 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 2224

Qy      1729 GGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTC 1788
             ||||::||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2225 GGTTMMGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTC 2284

Qy      1789 CCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGC 1848
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2285 CCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGC 2344

Qy      1849 CTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT 1908
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2345 CTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT 2404

Qy      1909 ATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG 1968
             ||||| |||||||||||||| |||||||||||||||| |||||||||||:||||||||||
Db      2405 ATTTGNCAAGAGCGACTTGG-CAAATACTCAGCGTAG-AAACTTCCAGCMCATTGGGGTG 2462

Qy      1969 GAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGG 2028
             :||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db      2463 RAGGGCCTGCCTCACTGGGTCCCAGCTCCCTGCTCCTGTTAGCCCCATGGGGCTGCCGGG 2522

Qy      2029 CTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTG 2088
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2523 CTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTG 2582

Qy      2089 CTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTC 2148
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2583 CTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTC 2642

Qy      2149 TAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGC 2208
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2643 TAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGC 2702

Qy      2209 CAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG 2268
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2703 CAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG 2762

Qy      2269 TTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAA 2328
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2763 TTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAA 2822

Qy      2329 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCT 2388
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2823 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCT 2882

Qy      2389 CTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTA 2446
             |||||||||||||||||||||||||||||||||||||||||||||||||||||  | |||
Db      2883 CTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGAAAGTTA 2942

Qy      2447 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 2506
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2943 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 3002

Qy      2507 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 2566
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3003 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 3062

Qy      2567 TTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGT 2626
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3063 TTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGT 3122

Qy      2627 AGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 2686
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3123 AGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 3182

Qy      2687 ATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCT 2746
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3183 ATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCT 3242

Qy      2747 TCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCA 2806
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3243 TCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCA 3302

Qy      2807 TCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGG 2866
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       3303 TCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGG 3362

Qy       2867 GTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGT 2926
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3363 GTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGT 3422

Qy       2927 TCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 2986
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3423 TCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 3482

Qy       2987 AAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCC 3046
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3483 AAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCC 3542

Qy       3047 TGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATC 3106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3543 TGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATC 3602

Qy       3107 TCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCC 3166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3603 TCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCC 3662

Qy       3167 TGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATG 3226
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3663 TGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATG 3722

Qy       3227 TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAA 3286
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3723 TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAA 3782

Qy       3287 TGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAA 3346
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3783 TGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAA 3842

Qy       3347 AAAA 3350
              ||||
Db       3843 AAAA 3846
```

```
RESULT 24
AAV71181
ID   AAV71181 standard; cDNA; 2152 BP.
XX
AC   AAV71181;
XX
DT   12-FEB-1999  (first entry)
XX
DE   Consensus sequence of the PS108 gene derived from overlapping clones.
XX
KW   PS108 gene; prostate disease; benign prostatic hyperplasia; BPH;
KW   prostatitis; prostatic intraepithelial neoplasia; PIN; cancer;
KW   drug screening; gene therapy; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             2..769
FT                   /*tag=  a
XX
PN   WO9850567-A1.
XX
PD   12-NOV-1998.
XX
PF   01-MAY-1998;   98WO-US08930.
XX
PR   02-MAY-1997;   97US-0850713.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Billing-Medel PA,  Cohen M,  Colpitts TL,  Friedman PN;
PI   Gordon J,  Granados EN,  Hodges SC,  Klass MR,  Kratochvil JD;
PI   Roberts-Rapp L,  Russell JC,  Stroupe SD;
XX
DR   WPI; 1999-034731/03.
DR   P-PSDB; AAW85068.
XX
PT   New isolated prostate-specific polynucleotides - used to develop
PT   products for the diagnosis and treatment of prostate diseases, e.g.
PT   benign hyperplasia, prostatic or prostate cancer
XX
PS   Claim 1; Fig 1A-E; 122pp; English.
XX
CC   The present sequence represents the consensus sequence for a PS108
CC   gene. The sequence is derived from overlapping clones AAV71166-79. The
CC   clone sequences are PS108 gene-specific. They are used in the method
CC   of the invention. The specification describes a method for detecting the
CC   presence of a target PS108 polynucleotide in a test sample. The method
CC   comprises contacting the test sample with at least 1 PS108-specific
CC   polynucleotide or complement, and detecting the presence of the target
CC   PS108 polynucleotide. The products can be used for detecting,
CC   diagnosing, staging, monitoring, prognosticating, in vivo imaging,
CC   preventing or treating, or determining predisposition to diseases or
CC   conditions of the prostate such as benign prostatic hyperplasia (BPH),
CC   prostatitis, prostatic intraepithelial neoplasia (PIN) and cancer. In
CC   particular the products can be used in drug screening and gene therapy.
XX
SQ   Sequence 2152 BP; 419 A; 622 C; 569 G; 542 T; 0 other;

  Query Match             62.7%;  Score 2136.4;  DB 20;  Length 2152;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 2149;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    1;

Qy       1177 GGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGA 1236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGA 60

Qy       1237 TGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTT 1296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTT 120

Qy       1297 CTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAG 1356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAG 180

Qy       1357 TGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGT 1416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGT 240

Qy       1417 GACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACAC 1476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACAC 300

Qy       1477 ACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACAC 1536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACAC 360

Qy       1537 TGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGG 1596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGG 420

Qy       1597 AGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACC 1656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACC 480

Qy       1657 CGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCAC 1716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCAC 540

Qy       1717 CGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGC 1776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGC 600

Qy       1777 CTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCA 1836
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCA 660
```

```
Qy       1837 GTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGC 1896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGC 720

Qy       1897 TACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAG 1956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAG 780

Qy       1957 CACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCAT 2016
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCAT 840

Qy       2017 GGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGC 2076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGC 900

Qy       2077 CACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCT 2136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCT 960

Qy       2137 CTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTT 2196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTT 1020

Qy       2197 ATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTAC 2256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTAC 1080

Qy       2257 CCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGG 2316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGG 1140

Qy       2317 GAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCG 2376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCG 1200

Qy       2377 TTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAAC 2436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAAC 1260

Qy       2437 ATATG--ACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCC 2494
              |||||  | |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCC 1320

Qy       2495 TCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTG 2554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTG 1380

Qy       2555 GCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTT 2614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTT 1440

Qy       2615 ATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGG 2674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGG 1500

Qy       2675 GTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCA 2734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCA 1560

Qy       2735 TTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGG 2794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGG 1620

Qy       2795 AAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAA 2854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAA 1680

Qy       2855 GGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTG 2914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTG 1740

Qy       2915 GCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGA 2974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGA 1800

Qy       2975 AGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCA 3034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 AGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCA 1860

Qy       3035 GCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAG 3094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAG 1920

Qy       3095 CCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTC 3154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTC 1980

Qy       3155 AGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCG 3214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 AGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCG 2040

Qy       3215 TTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCA 3274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCA 2100

Qy       3275 ATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTA 3326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 ATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTA 2152
```

```
RESULT 25
AAV71180
ID   AAV71180 standard; cDNA; 2143 BP.
XX
AC   AAV71180;
XX
DT   12-FEB-1999  (first entry)
XX
DE   Clone 1711346IH, the PS108 gene contig full length sequence.
XX
KW   PS108 gene; prostate disease; benign prostatic hyperplasia; BPH;
KW   prostatitis; prostatic intraepithelial neoplasia; PIN; cancer;
KW   drug screening; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO9850567-A1.
XX
PD   12-NOV-1998.
XX
PF   01-MAY-1998;   98WO-US08930.
XX
PR   02-MAY-1997;   97US-0850713.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Billing-Medel PA,  Cohen M,  Colpitts TL,  Friedman PN;
PI   Gordon J,  Granados EN,  Hodges SC,  Klass MR,  Kratochvil JD;
PI   Roberts-Rapp L,  Russell JC,  Stroupe SD;
XX
DR   WPI; 1999-034731/03.
XX
PT   New isolated prostate-specific polynucleotides - used to develop
PT   products for the diagnosis and treatment of prostate diseases, e.g.
PT   benign hyperplasia, prostatic or prostate cancer
XX
PS   Claim 1; Fig 1A-E; 122pp; English.
XX
CC   The present sequence represents the full length contig of the PS108
CC   gene, as represented by clone 1711346IH. This clone is the contig of
CC   overlapping clones AAV71166-79. The clone sequences are PS108
CC   gene-specific. They are used in the method of the invention. The
CC   specification describes a method for detecting the presence of a
CC   target PS108 polynucleotide in a test sample. The method comprises
```

```
CC   contacting the test sample with at least 1 PS108-specific polynucleotide
CC   or complement, and detecting the presence of the target PS108
CC   polynucleotide. The products can be used for detecting, diagnosing,
CC   staging, monitoring, prognosticating, in vivo imaging, preventing or
CC   treating, or determining predisposition to diseases or conditions of
CC   the prostate such as benign prostatic hyperplasia (BPH), prostatitis,
CC   prostatic intraepithelial neoplasia (PIN) and cancer. In particular the
CC   products can be used in drug screening and gene therapy.
XX
SQ   Sequence 2143 BP; 418 A; 621 C; 563 G; 541 T; 0 other;

  Query Match             62.0%;  Score 2114.8;  DB 20;  Length 2143;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 2139;  Conservative    0;  Mismatches    2;  Indels    3;  Gaps    2;

Qy       1185 ACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCG 1244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCG 60

Qy       1245 TTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGG 1304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGG 120

Qy       1305 TCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAG 1364
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAG 180

Qy       1365 CTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTT 1424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTT 240

Qy       1425 CAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCT 1484
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCT 300

Qy       1485 CCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTG 1544
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTG 360

Qy       1545 CTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCT 1604
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCT 420

Qy       1605 TCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCT 1664
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCT 480

Qy       1665 GCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCA 1724
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCA 540

Qy       1725 GGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGC 1784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGC 600

Qy       1785 TGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCA 1844
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCA 660

Qy       1845 CTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGG 1904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGG 720

Qy       1905 TAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGG 1964
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGG 780

Qy       1965 GGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGC 2024
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGC 840

Qy       2025 CGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGT 2084
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGT 900

Qy       2085 GCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAG 2144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAG 960

Qy       2145 TCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGG 2204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGG 1020

Qy       2205 AGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTC 2264
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db       1021 AGGCCAGAAGGGCTCCATGCACTGGAATGC-GGGACTCTGCAGGTGGATTACCCAGGCTC 1079

Qy       2265 AGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAA 2324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 AGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAA 1139

Qy       2325 TAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGT 2384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 TAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGT 1199

Qy       2385 AGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG--A 2442
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |
Db       1200 AGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGAAA 1259

Qy       2443 CTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCA 2502
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCA 1319

Qy       2503 CAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTG 2562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 CAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTG 1379

Qy       2563 GTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACA 2622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 GTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACA 1439

Qy       2623 AAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTG 2682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1440 AAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTG 1499

Qy       2683 GGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGA 2742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1500 GGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGA 1559

Qy       2743 ATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTA 2802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1560 ATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTA 1619

Qy       2803 CTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTA 2862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 CTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTA 1679

Qy       2863 GAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCC 2922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1680 GAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCC 1739

Qy       2923 TGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTG 2982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 TGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTG 1799

Qy       2983 CCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACA 3042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 CCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACA 1859

Qy       3043 ACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTC 3102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1860 ACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTC 1919

Qy       3103 CATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCAC 3162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCAC 1979
```

```
Qy       3163 CCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAAT 3222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1980 CCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAAT 2039

Qy       3223 AATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGT 3282
              |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db       2040 AATGTCGTCTTATTTTTTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGT 2099

Qy       3283 ATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTA 3326
              ||||||||||||||||||||||||||||||||||||||||||||
Db       2100 ATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTA 2143


RESULT 26
AAZ45677
ID   AAZ45677 standard; cDNA; 2462 BP.
XX
AC   AAZ45677;
XX
DT   06-APR-2000  (first entry)
XX
DE   cDNA sequence of a novel prostate cancer-associated gene.
XX
KW   Prostate cancer-associated gene; Incyte clone 1864683; bone cancer;
KW   cell proliferation; cancer; adrenal gland cancer; bladder cancer;
KW   prostate cancer; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             376..1071
FT                   /*tag=  a
FT                   /product= "prostate cancer-associated protein"
FT   sig_peptide     376..516
FT                   /*tag=  b
XX
PN   WO9967384-A2.
XX
PD   29-DEC-1999.
XX
PF   15-JUN-1999;   99WO-US13524.
XX
PR   22-JUN-1998;   98US-0102615.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Walker MG,  Volkmuth W,  Klingler TM,  Sprinzak EA;
XX
DR   WPI; 2000-126631/11.
DR   P-PSDB; AAY54369.
XX
PT   Identifying biomolecules for the diagnosis and treatment of diseases
PT   associated with cell-proliferation -
XX
PS   Claim 8; Page 47-48; 52pp; English.
XX
CC   The present sequence is derived from a prostate cancer-associated
CC   gene, and is represented by Incyte clone number 1864683. The sequence
CC   is used in the method of the invention. The specification
CC   describes a method for identifying biomolecules for the diagnosis or
CC   treatment of diseases associated with cell proliferation. The method
CC   comprises examining polynucleotides, consisting of prostate cancer-
CC   specific genes, and genes of unknown function, expressed in cDNA
CC   libraries. The patterns of both gene sets are compared to identify
CC   genes of unknown function with similar expression patterns to the
CC   prostate cancer-specific genes. The biomolecules identified by the
CC   method form pharmaceutical compositions useful for the diagnosis and
CC   treatment of diseases associated with cell proliferation. Such diseases
CC   include cancer of the adrenal gland, bladder and bone, but especially
CC   prostate cancer. The method may also be applied using other disease-
CC   specific genes. The prostate cancer-specific genes facilitate the
CC   diagnosis and treatment of cell proliferation disorders.
XX
SQ   Sequence 2462 BP; 483 A; 726 C; 632 G; 621 T; 0 other;

  Query Match             60.6%;  Score 2065.4;  DB 21;  Length 2462;
  Best Local Similarity   99.6%;  Pred. No. 7.6e-314;
  Matches 2092;  Conservative    0;  Mismatches    6;  Indels    3;  Gaps    2;

Qy       1237 TGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTT 1296
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 TGCAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTT 422

Qy       1297 CTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAG 1356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 CTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAG 482

Qy       1357 TGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGT 1416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 TGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGT 542

Qy       1417 GACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACAC 1476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 GACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACAC 602

Qy       1477 ACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACAC 1536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 ACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACAC 662

Qy       1537 TGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGG 1596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 TGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGG 722

Qy       1597 AGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACC 1656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 AGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACC 782

Qy       1657 CGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCAC 1716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 CGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCAC 842

Qy       1717 CGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGC 1776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 CGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGC 902

Qy       1777 CTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCA 1836
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        903 CTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCA 962

Qy       1837 GTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGC 1896
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db        963 GTCTGTCACTGCCTATATGGTGTCTGCCGCAGCGCTGGGTCTGGTCGCCATTTACTTTGC 1022

Qy       1897 TACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAG 1956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 TACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAG 1082

Qy       1957 CACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCAT 2016
              |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db       1083 CACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCTGCTCCTGTTAGCCCCAT 1142

Qy       2017 GGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGC 2076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1143 GGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGC 1202

Qy       2077 CACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCT 2136
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db       1203 CACCCTGTGCTGCTGAGGTGCGTA-CTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCT 1261

Qy       2137 CTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTT 2196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1262 CTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTT 1321

Qy       2197 ATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTAC 2256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1322 ATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTAC 1381

Qy       2257 CCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGG 2316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1382 CCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGG 1441

Qy       2317 GAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCG 2376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1442 GAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCG 1501

Qy       2377 TTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAAC 2436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1502 TTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAAC 1561

Qy       2437 ATATG--ACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCC 2494
              |||||  | |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1562 ATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCC 1621

Qy       2495 TCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTG 2554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1622 TCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTG 1681

Qy       2555 GCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTT 2614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1682 GCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTT 1741

Qy       2615 ATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGG 2674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1742 ATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGG 1801

Qy       2675 GTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCA 2734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1802 GTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCA 1861

Qy       2735 TTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGG 2794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1862 TTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGG 1921

Qy       2795 AAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAA 2854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1922 AAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAA 1981

Qy       2855 GGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTG 2914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1982 GGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTG 2041

Qy       2915 GCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGA 2974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2042 GCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGA 2101

Qy       2975 AGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCA 3034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2102 AGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCA 2161

Qy       3035 GCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAG 3094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2162 GCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAG 2221

Qy       3095 CCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTC 3154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2222 CCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTC 2281

Qy       3155 AGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCG 3214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2282 AGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCG 2341

Qy       3215 TTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCA 3274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2342 TTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCA 2401

Qy       3275 ATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAAAAAA 3334
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       2402 ATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTAAAAAAAAAA 2461

Qy       3335 A 3335
              |
Db       2462 A 2462


RESULT 27
AAC64928
ID   AAC64928 standard; DNA; 2133 BP.
XX
AC   AAC64928;
XX
DT   07-FEB-2001  (first entry)
XX
DE   Human prostate-related PS108 partial coding sequence SEQ ID NO: 16.
XX
KW   Human; prostate cancer; PS108; antibody; tumour; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   US6130043-A.
XX
PD   10-OCT-2000.
XX
PF   01-MAY-1998;   98US-0071710.
XX
PR   02-MAY-1997;   97US-0850713.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Cohen M,  Colpitts TL,  Friedman PN,  Gordon J,  Granados EN;
PI   Billing-Medel PA,  Klass MR,  Roberts-Rapp L,  Stroupe SD,  Yu H;
PI   Kratochvil JD,  Russell JC,  Hodges SC;
XX
DR   WPI; 2000-655655/63.
XX
PT   Methods for detecting target prostate-specific polynucleotides or
PT   diseases of the prostate (e.g. prostate cancer), comprising detecting
PT   the presence of any of PS108 nucleic acid sequences in a test sample -
XX
PS   Claim 1; Column 77-80; 55pp; English.
XX
CC   The present invention is related to a number of partial coding and
CC   protein sequences for the human prostate tissue protein PS108. These
CC   sequences can be used in the diagnosis and prognosis of prostate
CC   diseases, particularly prostate cancer. They can also be used to produce
CC   antibodies which can be used in treatment. The present sequence is one
CC   of the PS108 partial coding sequences.
XX
SQ   Sequence 2133 BP; 414 A; 618 C; 561 G; 540 T; 0 other;

  Query Match             56.2%;  Score 1915;  DB 21;  Length 2133;
  Best Local Similarity   98.9%;  Pred. No. 2.3e-290;
  Matches 2127;  Conservative    0;  Mismatches    5;  Indels   19;  Gaps   19;

Qy       1177 GGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGA 1236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGA 60

Qy       1237 TGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTT 1296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTT 120

Qy       1297 CTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAG 1356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAG 180

Qy       1357 TGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGT 1416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGT 240

Qy       1417 GACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACAC 1476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        241 GACAGCTTCAGCCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACAC 300

Qy       1477 ACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACAC 1536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACAC 360

Qy       1537 TGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGG 1596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGG 420

Qy       1597 AGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACC 1656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACC 480

Qy       1657 CGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCAC 1716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCAC 540

Qy       1717 CGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGC 1776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGC 600

Qy       1777 CTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCA 1836
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCA 660

Qy       1837 GTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGC 1896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGC 720

Qy       1897 TACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAG 1956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAG 780

Qy       1957 CACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCAT 2016
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCAT 840

Qy       2017 GGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGC 2076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGC 900

Qy       2077 CACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCT 2136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCT 960

Qy       2137 CTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTT 2196
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        961 CTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGAC-T 1019

Qy       2197 ATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTAC 2256
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 ATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTA- 1078

Qy       2257 CCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGG 2316
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1079 CCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTT-G 1137

Qy       2317 GAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCG 2376
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1138 GAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTC- 1196

Qy       2377 TTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTG-AA 2435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1197 TTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAAA 1256

Qy       2436 CATATGACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCT 2495
               ||   | ||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db       1257 TATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGT-CCCT 1315

Qy       2496 CAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGG 2555
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1316 CAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGT-G 1374

Qy       2556 CCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTA 2615
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       1375 CCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTA-TTA 1433

Qy       2616 TTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGG 2675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1434 TTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTT-GG 1492

Qy       2676 TAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCAT 2735
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1493 TAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATC-T 1551

Qy       2736 TGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGA 2795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1552 TGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTT-GA 1610

Qy       2796 AATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAG 2855
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1611 AATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTG-AG 1669

Qy       2856 GAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGG 2915
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1670 GAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTT-G 1728

Qy       2916 CCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAA 2975
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1729 CCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATG-A 1787

Qy       2976 GGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAG 3035
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1788 GGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACC-G 1846

Qy       3036 CTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGC 3095
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1847 CTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAA-C 1905

Qy       3096 CTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCA 3155
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1906 CTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACT-A 1964

Qy       3156 GGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGT 3215
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1965 GGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCC-T 2023

Qy       3216 TTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAA 3275
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       2024 TTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGC-A 2082

Qy       3276 TCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTA 3326
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2083 TCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTA 2133
```

```
RESULT 28
AAS07155
ID   AAS07155 standard; cDNA; 2133 BP.
XX
AC   AAS07155;
XX
DT   23-OCT-2001  (first entry)
XX
DE   Prostate gene PS108-specific cDNA (EST) consensus sequence.
XX
KW   Prostate; PS108; immunogen; drug screening; image localisation;
KW   diagnostic; therapeutic; prostate tissue disease; cancer; metastasis;
KW   expressed sequence tag; EST; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..799
FT                   /*tag=  a
```

```
FT                   /product= "Prostate-specific protein PS108"
XX
PN   US6252047-B1.
XX
PD   26-JUN-2001.
XX
PF   15-MAR-2000; 2000US-0525397.
XX
PR   01-MAY-1998;   98US-0071710.
PR   02-MAY-1997;   97US-0850713.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Billing-Medel PA,  Cohen M,  Colpitts TL,  Friedman PN,  Gordon J;
PI   Granados EN,  Hodges SC,  Klass MR,  Kratochvil JD,  Roberts-Rapp L;
PI   Russell JC,  Stroupe SD,  Yu H;
XX
DR   WPI; 2001-424488/45.
DR   P-PSDB; AAU04205.
XX
PT   Novel PS108 polypeptide useful in assays for detecting antibodies to
PT   prostate tissue, and as immunogens to produce PS108 antibodies  -
XX
PS   Example 1; Fig 1; 55pp; English.
XX
CC   The sequence represents the prostate gene PS108-specific expressed
CC   sequence tag (EST) consensus sequence. The sequence was produced from
CC   overlapping PS108 ESTs sequences to produce a full length consensus
CC   sequence. This sequence was then used to produce PS108 polypeptides which
CC   are useful in assays for detecting antibodies to prostate tissue, and as
CC   immunogens to produce antibodies. The polypeptide is useful for screening
CC   compounds which specifically bind to the polypeptide and for screening
CC   for drugs, compounds, or any other agent which can be used to treat
CC   diseases associated with PS108. The antibody is useful to detect, or for
CC   image localisation of, PS108 antigen in a patient, for detecting or
CC   diagnosing a disease or condition, as delivery agents for therapeutic
CC   agents as well as for diagnostic tests and for screening for diseases or
CC   conditions associated with PS108, especially cancer. The antibody is also
CC   useful for generating chimeric antibodies for therapeutic use, for
CC   inhibiting the biological activity of PS108, in therapy (for e.g. to
CC   treat prostate tissue disease including prostate cancer and its
CC   metastases), and to detect the presence of any polypeptide in a test
CC   sample which shares one or more antigenic determinants with the PS108
CC   polypeptide.
XX
SQ   Sequence 2133 BP; 414 A; 618 C; 561 G; 540 T; 0 other;

  Query Match             56.2%;  Score 1915;  DB 22;  Length 2133;
  Best Local Similarity   98.9%;  Pred. No. 2.3e-290;
  Matches 2127;  Conservative    0;  Mismatches    5;  Indels   19;  Gaps   19;

Qy       1177 GGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGA 1236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGA 60

Qy       1237 TGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTT 1296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTT 120

Qy       1297 CTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAG 1356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAG 180

Qy       1357 TGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGT 1416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGT 240

Qy       1417 GACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACAC 1476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACAC 300

Qy       1477 ACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACAC 1536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACAC 360

Qy       1537 TGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGG 1596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGG 420

Qy       1597 AGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACC 1656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACC 480

Qy       1657 CGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCAC 1716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCAC 540

Qy       1717 CGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGC 1776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGC 600

Qy       1777 CTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCA 1836
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCA 660

Qy       1837 GTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGC 1896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGC 720

Qy       1897 TACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAG 1956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAG 780

Qy       1957 CACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCAT 2016
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCAT 840

Qy       2017 GGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGC 2076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGC 900

Qy       2077 CACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCT 2136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCT 960

Qy       2137 CTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTT 2196
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        961 CTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGAC-T 1019

Qy       2197 ATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTAC 2256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db       1020 ATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTA- 1078

Qy       2257 CCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGG 2316
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1079 CCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTT-G 1137

Qy       2317 GAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCG 2376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db       1138 GAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTC- 1196

Qy       2377 TTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTG-AA 2435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1197 TTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAAA 1256

Qy       2436 CATATGACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCT 2495
               ||   | ||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db       1257 TATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGT-CCCT 1315

Qy       2496 CAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGG 2555
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1316 CAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGT-G 1374

Qy       2556 CCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTA 2615
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       1375 CCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTA-TTA 1433
```

```
Qy       2616 TTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGG 2675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1434 TTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTT-GG 1492

Qy       2676 TAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCAT 2735
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1493 TAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATC-T 1551

Qy       2736 TGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGA 2795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1552 TGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTT-GA 1610

Qy       2796 AATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAG 2855
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1611 AATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTG-AG 1669

Qy       2856 GAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGG 2915
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1670 GAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTT-G 1728

Qy       2916 CCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAA 2975
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1729 CCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATG-A 1787

Qy       2976 GGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAG 3035
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1788 GGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACC-G 1846

Qy       3036 CTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGC 3095
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1847 CTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAA-C 1905

Qy       3096 CTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCA 3155
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1906 CTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACT-A 1964

Qy       3156 GGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGT 3215
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1965 GGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCC-T 2023

Qy       3216 TTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAA 3275
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       2024 TTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGC-A 2082

Qy       3276 TCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTA 3326
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2083 TCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTA 2133
```

RESULT 29
AAC64927

```
ID   AAC64927 standard; DNA; 2124 BP.
XX
AC   AAC64927;
XX
DT   07-FEB-2001  (first entry)
XX
DE   Human prostate-related PS108 partial coding sequence SEQ ID NO: 15.
XX
KW   Human; prostate cancer; PS108; antibody; tumour; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   US6130043-A.
XX
PD   10-OCT-2000.
XX
PF   01-MAY-1998;   98US-0071710.
XX
PR   02-MAY-1997;   97US-0850713.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Cohen M,  Colpitts TL,  Friedman PN,  Gordon J,  Granados EN;
PI   Billing-Medel PA,  Klass MR,  Roberts-Rapp L,  Stroupe SD,  Yu H;
PI   Kratochvil JD,  Russell JC,  Hodges SC;
XX
DR   WPI; 2000-655655/63.
XX
PT   Methods for detecting target prostate-specific polynucleotides or
PT   diseases of the prostate (e.g. prostate cancer), comprising detecting
PT   the presence of any of PS108 nucleic acid sequences in a test sample -
XX
PS   Claim 1; Column 75-78; 55pp; English.
XX
CC   The present invention is related to a number of partial coding and
CC   protein sequences for the human prostate tissue protein PS108. These
CC   sequences can be used in the diagnosis and prognosis of prostate
CC   diseases, particularly prostate cancer. They can also be used to produce
CC   antibodies which can be used in treatment. The present sequence is one
CC   of the PS108 partial coding sequences.
XX
SQ   Sequence 2124 BP; 411 A; 616 C; 558 G; 539 T; 0 other;

  Query Match             55.7%;  Score 1899.8;  DB 21;  Length 2124;
  Best Local Similarity   99.0%;  Pred. No. 5.5e-288;
  Matches 2121;  Conservative    0;  Mismatches    2;  Indels   20;  Gaps   20;

Qy       1185 ACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCG 1244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCG 60

Qy       1245 TTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGG 1304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGG 120

Qy       1305 TCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAG 1364
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAG 180

Qy       1365 CTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTT 1424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTT 240

Qy       1425 CAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCT 1484
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCT 300

Qy       1485 CCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTG 1544
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTG 360

Qy       1545 CTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCT 1604
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCT 420

Qy       1605 TCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCT 1664
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCT 480

Qy       1665 GCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCA 1724
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCA 540

Qy       1725 GGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGC 1784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGC 600

Qy       1785 TGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCA 1844
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCA 660

Qy       1845 CTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGG 1904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGG 720
```

```
Qy      1905 TAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGG 1964
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 TAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGG 780

Qy      1965 GGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGC 2024
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 GGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGC 840

Qy      2025 CGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGT 2084
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 CGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGT 900

Qy      2085 GCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAG 2144
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 GCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAG 960

Qy      2145 TCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGG 2204
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||
Db       961 TCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACA-GG 1019

Qy      2205 AGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTC 2264
             |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db      1020 AGGCCAGAAGGGCTCCATGCACTGGAATGC-GGGACTCTGCAGGTGGATTACCCAGGCTC 1078

Qy      2265 AGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAA 2324
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1079 -GGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAA 1137

Qy      2325 TAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGT 2384
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1138 -AAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGT 1196

Qy      2385 AGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG-AC 2443
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||  |
Db      1197 -GCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGAAA 1255

Qy      2444 TTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCAC 2503
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db      1256 TTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCA- 1314

Qy      2504 AGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGG 2563
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      1315 AGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTT-G 1373

Qy      2564 TCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAA 2623
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      1374 TCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAAC-A 1432

Qy      2624 AGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGG 2683
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      1433 AGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGT-G 1491

Qy      2684 GGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAA 2743
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      1492 GGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAG-A 1550

Qy      2744 TCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTAC 2803
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db      1551 TCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTA- 1609

Qy      2804 TCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAG 2863
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db      1610 TCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTA- 1668

Qy      2864 AGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCT 2923
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db      1669 AGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCC- 1727

Qy      2924 GGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGC 2983
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db      1728 GGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTG- 1786

Qy      2984 CCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAA 3043
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      1787 CCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCAC-A 1845

Qy      3044 CCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCC 3103
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      1846 CCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGT-C 1904

Qy      3104 ATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACC 3163
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      1905 ATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCA-C 1963

Qy      3164 CCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATA 3223
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db      1964 CCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAAT- 2022

Qy      3224 ATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTA 3283
             |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||| 
Db      2023 ATGTCGTCTTATTTTTTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGT- 2081

Qy      3284 TAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTA 3326
             |||||||||||||||||||||||||||||||||||||||||||
Db      2082 TAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTA 2124
```

```
RESULT 30
AAS07601
ID   AAS07601 standard; cDNA; 2124 BP.
XX
AC   AAS07601;
XX
DT   23-OCT-2001  (first entry)
XX
DE   Prostate gene PS108-specific cDNA clone 1711346IH.
XX
KW   Prostate; PS108; immunogen; drug screening; image localisation;
KW   diagnostic; therapeutic; prostate tissue disease; cancer; metastasis;
KW   expressed sequence tag; EST; ss.
XX
OS   Homo sapiens.
XX
PN   US6252047-B1.
XX
PD   26-JUN-2001.
XX
PF   15-MAR-2000; 2000US-0525397.
XX
PR   01-MAY-1998;   98US-0071710.
PR   02-MAY-1997;   97US-0850713.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Billing-Medel PA,  Cohen M,  Colpitts TL,  Friedman PN,  Gordon J;
PI   Granados EN,  Hodges SC,  Klass MR,  Kratochvil JD,  Roberts-Rapp L;
PI   Russell JC,  Stroupe SD,  Yu H;
XX
DR   WPI; 2001-424488/45.
XX
PT   Novel PS108 polypeptide useful in assays for detecting antibodies to
PT   prostate tissue, and as immunogens to produce PS108 antibodies  -
XX
PS   Example 1; Fig 1; 55pp; English.
XX
CC   The sequence represents the coding sequence of prostate gene PS108-
CC   specific expressed sequence tag (EST) cDNA clone 1711346IH. The sequence
CC   was used along with other overlapping cDNA clones to produce a full
CC   length consensus sequence (see AAS07155). This sequence could then be
CC   used to produce PS108 polypeptides which are useful in assays for
CC   detecting antibodies to prostate tissue, and as immunogens to produce
CC   antibodies. The polypeptide is useful for screening compounds which
CC   specifically bind to the polypeptide and for screening for drugs,
CC   compounds, or any other agent which can be used to treat diseases
CC   associated with PS108. The antibody is useful to detect, or for image
```

```
CC   localisation of PS108 antigen in a patient, for detecting or diagnosing a
CC   disease or condition, as delivery agents for therapeutic agents as well
CC   as for diagnostic tests and for screening for diseases or conditions
CC   associated with PS108, especially cancer. The antibody is also useful for
CC   generating chimeric antibodies for therapeutic use, for inhibiting the
CC   biological activity of PS108, in therapy (for e.g. to treat prostate
CC   tissue disease including prostate cancer and its metastases), and to
CC   detect the presence of any polypeptide in a test sample which shares one
CC   or more antigenic determinants with the PS108 polypeptide.
XX
SQ   Sequence 2124 BP; 411 A; 616 C; 558 G; 539 T; 0 other;

  Query Match             55.7%;  Score 1899.8;  DB 22;  Length 2124;
  Best Local Similarity   99.0%;  Pred. No. 5.5e-288;
  Matches 2121;  Conservative    0;  Mismatches    2;  Indels   20;  Gaps   20;

Qy       1185 ACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCG 1244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCG 60

Qy       1245 TTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGG 1304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGG 120

Qy       1305 TCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAG 1364
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAG 180

Qy       1365 CTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTT 1424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTT 240

Qy       1425 CAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCT 1484
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCT 300

Qy       1485 CCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTG 1544
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTG 360

Qy       1545 CTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCT 1604
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCT 420

Qy       1605 TCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCT 1664
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCT 480

Qy       1665 GCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCA 1724
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCA 540

Qy       1725 GGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGC 1784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGC 600

Qy       1785 TGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCA 1844
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCA 660

Qy       1845 CTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGG 1904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGG 720

Qy       1905 TAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGG 1964
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGG 780

Qy       1965 GGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGC 2024
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGC 840

Qy       2025 CGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGT 2084
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGT 900

Qy       2085 GCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAG 2144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAG 960

Qy       2145 TCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGG 2204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        961 TCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACA-GG 1019

Qy       2205 AGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTC 2264
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db       1020 AGGCCAGAAGGGCTCCATGCACTGGAATGC-GGGACTCTGCAGGTGGATTACCCAGGCTC 1078

Qy       2265 AGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTGGGAGCTGAA 2324
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 -GGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTGGGAGCTGAA 1137

Qy       2325 TAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGT 2384
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1138 -AAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGT 1196

Qy       2385 AGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG-AC 2443
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||| | 
Db       1197 -GCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGAAA 1255

Qy       2444 TTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCAC 2503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db       1256 TTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCA- 1314

Qy       2504 AGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGG 2563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| | 
Db       1315 AGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTT-G 1373

Qy       2564 TCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAA 2623
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1374 TCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAAC-A 1432

Qy       2624 AGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGG 2683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1433 AGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGT-G 1491

Qy       2684 GGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAA 2743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1492 GGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAG-A 1550

Qy       2744 TCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTAC 2803
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db       1551 TCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTA- 1609

Qy       2804 TCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAG 2863
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db       1610 TCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTA- 1668

Qy       2864 AGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCT 2923
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db       1669 AGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCC- 1727

Qy       2924 GGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGC 2983
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db       1728 GGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTG- 1786

Qy       2984 CCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAA 3043
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1787 CCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCAC-A 1845

Qy       3044 CCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCC 3103
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1846 CCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGT-C 1904

Qy       3104 ATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACC 3163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
```

```
Db       1905 ATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCA-C 1963

Qy       3164 CCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATA 3223
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1964 CCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAAT- 2022

Qy       3224 ATGTCGTCTTATTTATTTAGCGGGGTGAATATTTATACTGTAAGTGAGCAATCAGAGTA 3283
              ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db       2023 ATGTCGTCTTATTTTTTTAGCGGGGTGAATATTTATACTGTAAGTGAGCAATCAGAGT- 2081

Qy       3284 TAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTA 3326
              |||||||||||||||||||||||||||||||||||||||||||
Db       2082 TAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTA 2124


RESULT 31
ABN81320
ID   ABN81320 standard; cDNA; 3663 BP.
XX
AC   ABN81320;
XX
DT   30-AUG-2002  (first entry)
XX
DE   Human mast cell related gene MC14 SEQ ID NO 3.
XX
KW   Human; mast cell; MC; antiallergic; antiinflammatory; antiasthmatic;
KW   vasotropic; dermatological; allergy; hypersensitivity; rhinitis; asthma;
KW   gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             294..746
FT                   /*tag=  a
FT                   /product= "MC14-1"
FT   CDS             1238..2218
FT                   /*tag=  a
FT                   /product= "MC14-2"
XX
PN   WO200246389-A2.
XX
PD   13-JUN-2002.
XX
PF   07-DEC-2001; 2001WO-US46180.
XX
PR   08-DEC-2000; 2000US-251835P.
PR   14-MAR-2001; 2001US-275479P.
PR   28-MAR-2001; 2001US-279115P.
PR   02-APR-2001; 2001US-280143P.
XX
PA   (UNIO ) UCB SA.
XX
PI   Nocka K,  Pirozzi G,  Einstein R;
XX
DR   WPI; 2002-508560/54.
DR   P-PSDB; ABB77570, ABB77571.
XX
PT   Novel isolated nucleic acids that are differentially expressed in mast
PT   cells in patients with allergic hypersensitivity, encoding proteins
PT   associated with mast cell regranulation and allergic hypersensitivity
PT   -
XX
PS   Claim 1; Page 98-102; 119pp; English.
XX
CC   The invention relates to isolated nucleic acid (ABN81319-ABN81324),
CC   corresponding to genes differentially expressed in mast cells following
CC   activation or in patients with allergic hypersensitivity disease, (I)
CC   that encodes proteins (ABB77569-ABB77575) (II) or a protein fragment of
CC   (II) if at least 6 amino acids. (II) is useful for identifying binding
CC   partners. (I) or (II) is useful for diagnosing or treating a disease
CC   state (e.g. allergic hypersensitivity, seasonal rhinitis, asthma,
CC   urticaria or atopic dermatitis or mastocytosis) in a subject which
CC   involves determining the level of expression of (I) or (II). A computer
CC   system, comprising a database containing information identifying the
CC   expression level in a tissue or at least one mast cell of (I), is useful
CC   for presenting information to identify the relative expression level of
CC   (I). (II) is used as a marker to detect, diagnose or identify an allergic
CC   response in a patient. The protein can also serve as a target that
CC   modulate gene expression or activity and as an antigen to raise
CC   polyclonal or monoclonal antibodies. (II) is useful for identifying
CC   agents that modulate expression of the protein or agents, such as
CC   agonists or antagonists. The agonists or antagonists are useful for
CC   modulating biological activity and function of (II) and thus are useful
CC   for alleviating disease conditions such as allergic hypersensitivity,
CC   seasonal rhinitis, asthma, urticaria, atopic dermatitis or mastocytosis.
XX
SQ   Sequence 3663 BP; 740 A; 1058 C; 985 G; 880 T; 0 other;

  Query Match             54.4%;  Score 1854.8;  DB 24;  Length 3663;
  Best Local Similarity   73.0%;  Pred. No. 6.2e-281;
  Matches 2667;  Conservative    0;  Mismatches  682;  Indels  303;  Gaps   10;

Qy          2 GGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAG 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 GGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAG 71

Qy         62 TGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGA 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 TGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGA 131

Qy        122 AGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTG 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 AGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTG 191

Qy        182 GCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGG 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 GCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGG 251

Qy        242 AGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGG 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 AGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGG 311

Qy        302 GTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTT 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 GTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTT 371

Qy        362 GGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTG 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 GGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTG 431

Qy        422 GGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGG------------------------- 456
              |||||||||||||||||||||||||||||||||||
Db        432 GGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGTGAGTCACTACATCCTCCTTCCTTC 491

Qy        457 -------------------------CATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGC 491
                                       |||  |||  |    |||  |||  |  | ||
Db        492 CTGTTCCAGATACATGCCACCTGGCATGTGGGACAGGAGTACCTCTGCCCTGGGAGCTGC 551

Qy        492 TCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCT 551
               |  |||   ||   ||   |   |   |    |||   |    |   |  ||||
Db        552 TTGGAGGGAGAGGTGGTCTGCTGGGAAGGCATTGCTGGGCAGGAGGGTGACCCTGGGCTG 611

Qy        552 GGG---------------------------------------CACTGTCCTTGGGCAT 570
               ||                                        || |  |||  ||
Db        612 AGGGGGCACACCAAGAGAAAGAAGAGAATACCAAGGACATACCCAGTCACCTCTGGATC 671

Qy        571 CCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGA 630
              |||| |    | |     |   |   |   |  |  |     || |   ||| ||    |   |
Db        672 CCTGGTCCTGCACAGAGCCTGGCTCATAGGAGACACTGGAGAAATGCTCCTAACCTTTGG 731

Qy        631 ---------------------------------------TCCCAGGCCCCTGGAGCTG 649
                                                      |  ||  ||  ||    |
Db        732 CTAGCCCTTTTATAATTTATAGCGATTATCTCATTTAATGCTTACAACCACCATTTGAGG 791

Qy        650 GCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCA 709
                |     |  | | | |    |  | |   |   ||  |  ||  ||||  |  |  |  | |
```

```
Db     792 TGATCCATTTTACAGAGAAGGAAGCAGAGGCTTTTAAGAGGTTAGGTAAGTCTTAGCCAA 851

Qy     710 CTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCT 769
Db     852 AGCCAAATAGCAGCTGAACAGTAGAGCTGGGACTCCATCAAGGTCTCCCAGCCGGAGCTT 911

Qy     770 GTCTATGCCTTCATGATCAGTCTTGG---------------------------------- 795
Db     912 GCTCCTACCCCTAGGACAAGGGGTGGACTCCTGACTCTGCAGATAAATTCTACAAAAGCC 971

Qy     796 ----------------GGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCA 839
Db     972 ACAGAAGGCAAGTAGTAACCATTGTGTGACAACCCCTCACCCCCAGGAAGAGGGGCCCCT 1031

Qy     840 GTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCA 899
Db    1032 GTGAGGATTGCAGGCTCTGGAGTCACACTGCTTGTTGAAACGCTGCCTCTTACCCTCCCT 1091

Qy     900 TCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCA 959
Db    1092 AGGTCTGCGCCTTTGAATAAGTATCACTTCTTAGTTGCTGCATGCCTCAGTTTGTCCATC 1151

Qy     960 ------------------------------------------------------------ 959
Db    1152 TGAAAATGGGGGCATCTGTAATGCCTGTGTTATGAGGAGTAAATTACAGCATCCCTGTGA 1211

Qy     960 ----------------------------------CCGAGCCAGCAGAAGGGCTGTCGGCC 985
Db    1212 AGACGTAGCACAGTGTCGAGTACGGAATGTTATTTCCATCCTTCTCACGGAGCTTGGTTC 1271

Qy     986 CCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGC 1045
Db    1272 CCCTTCCCCTTGCCCTTTACTTGTCCCAGCCATTGACTCATACTACTTCCCTTCTTGCAG 1331

Qy    1046 GCCCTG--------------CTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCA 1091
Db    1332 GCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACT 1391

Qy    1092 CCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGC 1151
Db    1392 GGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGC 1451

Qy    1152 TGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGG 1211
Db    1452 TGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCA 1511

Qy    1212 GCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCC 1271
Db    1512 GGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGG 1571

Qy    1272 TGCAGTGCGC-CATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTC 1330
Db    1572 TGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTC 1631

Qy    1331 GGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACA 1390
Db    1632 GCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCC 1691

Qy    1391 TGCCTGTCCCACAGTGTGGCCGTGGTGA------CAGCTTCAGCCGCCCTCACCGGGTTC 1444
Db    1692 TGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGT 1751

Qy    1445 ACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAG 1504
Db    1752 GCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGG 1811

Qy    1505 CAGGTGTTCCTGCCC----AAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
Db    1812 CTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGT 1871

Qy    1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
Db    1872 CGCCCCACTGCTGTCCATGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1931

Qy    1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
Db    1932 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1991

Qy    1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
Db    1992 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 2051

Qy    1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
Db    2052 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 2111

Qy    1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
Db    2112 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 2171

Qy    1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
Db    2172 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 2231

Qy    1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
Db    2232 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 2291

Qy    1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
Db    2292 CACTGGGTCCCAGCTCCCTGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2351

Qy    2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
Db    2352 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2411

Qy    2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
Db    2412 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2471

Qy    2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
Db    2472 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2531

Qy    2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
Db    2532 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2591

Qy    2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
Db    2592 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2651

Qy    2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
Db    2652 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2711

Qy    2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAA 2458
Db    2712 TTTCTAGGATGAAACACTCCACCATGGGATTTGAACATATGAAAGTTATTTGTAGGGGAA 2771

Qy    2459 GAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTG 2518
Db    2772 GAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTG 2831

Qy    2519 CTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCAT 2578
Db    2832 CTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCAT 2891

Qy    2579 CACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCC 2638
Db    2892 CACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCC 2951

Qy    2639 ATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAAT 2698
Db    2952 ATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAAT 3011
```

```
Qy        2699 CAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGG 2758
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3012 CAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGG 3071

Qy        2759 GTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATA 2818
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3072 GTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATA 3131

Qy        2819 ATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAG 2878
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3132 ATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAG 3191

Qy        2879 GTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTC 2938
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3192 GTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTC 3251

Qy        2939 CACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTA 2998
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3252 CACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTA 3311

Qy        2999 CCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTA 3058
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3312 CCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTA 3371

Qy        3059 CTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGA 3118
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3372 CTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGA 3431

Qy        3119 GTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAA 3178
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3432 GTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAA 3491

Qy        3179 GGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTA 3238
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3492 GGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTA 3551

Qy        3239 TTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGA 3298
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3552 TTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGA 3611

Qy        3299 CAAAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAA 3350
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3612 CAAAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAA 3663

RESULT 32
AAS64041
ID   AAS64041 standard; cDNA; 6976 BP.
XX
AC   AAS64041;
XX
DT   29-JAN-2002  (first entry)
XX
DE   Human prosate cDNA P553S splice variant #4.
XX
KW   Human; prostate cancer; ss; cytostatic; immunostimulant; tumour.
XX
OS   Homo sapiens.
XX
PN   WO200173032-A2.
XX
PD   04-OCT-2001.
XX
PF   27-MAR-2001; 2001WO-US09919.
XX
PR   27-MAR-2000; 2000US-0536857.
PR   09-MAY-2000; 2000US-0568100.
PR   12-MAY-2000; 2000US-0570737.
PR   13-JUN-2000; 2000US-0593793.
PR   27-JUN-2000; 2000US-0605783.
PR   10-AUG-2000; 2000US-0636215.
PR   29-AUG-2000; 2000US-0651236.
PR   06-SEP-2000; 2000US-0657279.
PR   02-OCT-2000; 2000US-0679426.
PR   10-OCT-2000; 2000US-0685166.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,   Dillon DC,   Mitcham JL,   Harlocker SL,   Jiang Y,   Kalos MD;
PI   Fanger GR,   Retter MW,   Stolk JA,   Day CH,   Vedvick TS,   Carter D;
PI   Li SX,   Wang A,   Skeiky YAW,   Hepler WT,   Henderson RA;
XX
DR   WPI; 2001-639232/73.
DR   P-PSDB; AAU69873.
XX
PT   New human prostate-specific polypeptides and polynucleotides useful for
PT   the diagnosis and treatment of cancer, especially prostate cancer -
XX
PS   Claim 1; Page 484-486; 579pp; English.
XX
CC   The invention relates to isolated prostate-specific
CC   polynucleotides, polypeptides, fusion proteins of the polypeptides,
CC   antibodies raised against the polypeptides (or antigenic epitopes
CC   derived from them) and antigen-presenting cells expressing the
CC   polypeptides. The antibodies are useful for detecting the presence of
CC   cancer, especially prostate cancer. The polypeptides, polynucleotides and
CC   the antigen-presenting cells are useful for stimulating and/or expanding
CC   T cells specific for a tumour protein, and for inhibiting the development
CC   of cancer especially prostate cancer. Compositions comprising the
CC   polynucleotide and/or polypeptide are useful for stimulating an immune
CC   response, and for treating cancer. The oligonucleotide is useful for
CC   detecting cancer. The present sequence is a prostate specific
CC   polynucleotide of the invention.
XX
SQ   Sequence 6976 BP; 1312 A; 1996 C; 1961 G; 1706 T; 1 other;

  Query Match             53.2%;  Score 1815.8;  DB 22;  Length 6976;
  Best Local Similarity   98.7%;  Pred. No. 8.2e-275;
  Matches 1841;  Conservative    0;  Mismatches   22;  Indels    2;  Gaps    1;

Qy        1470 CCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAG 1529
                |||  | ||  ||    ||| |  ||    |     |||||||||||||||||||||||
Db        5112 CCTGTCCTCTTCCCCTTTTCTTCACCCCTCTGCCTTAGGTGTTCCTGCCCAAATACCGAG 5171

Qy        1530 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTA 1589
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5172 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTA 5231

Qy        1590 AGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCAC 1649
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5232 AGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCAC 5291

Qy        1650 CTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTG 1709
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5292 CTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTG 5351

Qy        1710 AGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGG 1769
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5352 AGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGG 5411

Qy        1770 ATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGC 1829
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5412 ATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGC 5471

Qy        1830 TCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTT 1889
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5472 TCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTT 5531

Qy        1890 ACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAA 1949
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5532 ACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAA 5591

Qy        1950 CTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTA 2009
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5592 CTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTA 5651
```

```
Qy      2010 GCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCT 2069
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5652 GCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCT 5711

Qy      2070 CTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT 2129
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5712 CTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT 5771

Qy      2130 CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTC 2189
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5772 CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTC 5831

Qy      2190 TGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT 2249
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5832 TGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT 5891

Qy      2250 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG 2309
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5892 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG 5951

Qy      2310 TTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGC 2369
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5952 TTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGC 6011

Qy      2370 AGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGA 2429
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6012 AGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGA 6071

Qy      2430 TTTGAACATATGAC--TTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCA 2487
             |||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db      6072 TTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCA 6131

Qy      2488 GGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCA 2547
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6132 GGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCA 6191

Qy      2548 GGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA 2607
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6192 GGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA 6251

Qy      2608 CTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAA 2667
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6252 CTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAA 6311

Qy      2668 TATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGG 2727
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6312 TATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGG 6371

Qy      2728 CTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGG 2787
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6372 CTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGG 6431

Qy      2788 ACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGG 2847
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6432 ACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGG 6491

Qy      2848 TGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCT 2907
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6492 TGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCT 6551

Qy      2908 TCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGG 2967
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6552 TCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGG 6611

Qy      2968 CTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTC 3027
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6612 CTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTC 6671

Qy      3028 CCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 3087
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6672 CCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 6731

Qy      3088 TCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACA 3147
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6732 TCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACA 6791

Qy      3148 GAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTA 3207
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6792 GAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTA 6851

Qy      3208 AGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 3267
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6852 AGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 6911

Qy      3268 GTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAA 3327
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6912 GTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAA 6971

Qy      3328 AAAAA 3332
             |||||
Db      6972 AAAAA 6976
```

```
RESULT 33
AAH93869
ID   AAH93869 standard; cDNA; 6976 BP.
XX
AC   AAH93869;
XX
DT   04-OCT-2001  (first entry)
XX
DE   P553S cDNA splice variant P553S-6.
XX
KW   Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW   cytostatic; gene therapy; metastasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200151633-A2.
XX
PD   19-JUL-2001.
XX
PF   16-JAN-2001; 2001WO-US01574.
XX
PR   14-JAN-2000; 2000US-0483672.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Reed SG;
PI   Kalos MD,  Fanger GR,  Day CH,  Retter MW,  Stolk JA,  Skeiky YAW;
PI   Wang A,  Meagher MJ;
XX
DR   WPI; 2001-425873/45.
XX
PT   New polynucleotide encoding a prostate-specific protein, for
PT   diagnosing, monitoring and treating prostate cancer in a patient and
PT   for use in vaccines -
XX
PS   Claim 1; Page 461-463; 543pp; English.
XX
CC   The present invention describes polynucleotide sequences (I) which encode
CC   prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC   and can be used in vaccine production and gene therapy. (I), (II),
CC   antibodies to (II), fusion proteins comprising (II), and isolated
CC   T cells prepared using (I) or (II) are used treat cancer in a patient.
CC   (I) and the antibodies are also used in the detection of cancer in a
CC   patient. The cancer that is diagnosed or treated is particularly
CC   prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
CC   (I) can be used for monitoring the progression of cancer in a patient.
CC   (I) and (II) can also be used to improve diagnostic and therapeutic
CC   methods for prostate cancer. They can indicate the level of metastasis
CC   as well as the prostate volume. AAH93357 to AAH93944 and AAM01115 to
CC   AAM01318 represent polynucleotide and amino acid sequences used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 6976 BP; 1312 A; 1996 C; 1961 G; 1706 T; 1 other;
```

```
Query Match             53.2%;  Score 1815.8;  DB 22;  Length 6976;
Best Local Similarity   98.7%;  Pred. No. 8.2e-275;
Matches 1841;  Conservative    0;  Mismatches   22;  Indels    2;  Gaps    1;

Qy       1470 CCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAG 1529
              |||   | ||  ||    ||| |  ||    |    ||||||||||||||||||||||||
Db       5112 CCTGTCCTCTTCCCCTTTTCTTCACCCCTCTGCCTTAGGTGTTCCTGCCCAAATACCGAG 5171

Qy       1530 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTA 1589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5172 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTA 5231

Qy       1590 AGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCAC 1649
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5232 AGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCAC 5291

Qy       1650 CTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTG 1709
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5292 CTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTG 5351

Qy       1710 AGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGG 1769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5352 AGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGG 5411

Qy       1770 ATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGC 1829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5412 ATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGC 5471

Qy       1830 TCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTT 1889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5472 TCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTT 5531

Qy       1890 ACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAA 1949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5532 ACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAA 5591

Qy       1950 CTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTA 2009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5592 CTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTA 5651

Qy       2010 GCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCT 2069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5652 GCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCT 5711

Qy       2070 CTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT 2129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5712 CTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT 5771

Qy       2130 CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTC 2189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5772 CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTC 5831

Qy       2190 TGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT 2249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5832 TGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT 5891

Qy       2250 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG 2309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5892 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG 5951

Qy       2310 TTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGC 2369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5952 TTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGC 6011

Qy       2370 AGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGA 2429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6012 AGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGA 6071

Qy       2430 TTTGAACATATGAC--TTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCA 2487
              |||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db       6072 TTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCA 6131

Qy       2488 GGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCA 2547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6132 GGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCA 6191

Qy       2548 GGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA 2607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6192 GGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA 6251

Qy       2608 CTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAA 2667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6252 CTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAA 6311

Qy       2668 TATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGG 2727
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6312 TATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGG 6371

Qy       2728 CTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGG 2787
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6372 CTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGG 6431

Qy       2788 ACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGG 2847
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6432 ACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGG 6491

Qy       2848 TGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCT 2907
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6492 TGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCT 6551

Qy       2908 TCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGG 2967
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6552 TCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGG 6611

Qy       2968 CTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTC 3027
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6612 CTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTC 6671

Qy       3028 CCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 3087
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6672 CCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 6731

Qy       3088 TCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACA 3147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6732 TCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACA 6791

Qy       3148 GAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTA 3207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6792 GAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTA 6851

Qy       3208 AGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 3267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6852 AGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 6911

Qy       3268 GTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAA 3327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6912 GTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAA 6971

Qy       3328 AAAAA 3332
              |||||
Db       6972 AAAAA 6976


RESULT 34
ABL95412
ID   ABL95412 standard; cDNA; 6976 BP.
XX
AC   ABL95412;
XX
DT   19-JUL-2002  (first entry)
XX
DE   Human P553S splice variant SEQ ID NO 705.
XX
KW   Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
KW   gene therapy; gene; ss.
```

```
XX
OS   Homo sapiens.
XX
PN   US2002022248-A1.
XX
PD   21-FEB-2002.
XX
PF   12-JAN-2001; 2001US-0759143.
XX
PR   25-FEB-1997;   97US-0806099.
PR   01-AUG-1997;   97US-0904804.
PR   09-FEB-1998;   98US-0020956.
PR   25-FEB-1998;   98US-0030607.
PR   14-JUL-1998;   98US-0115453.
PR   23-SEP-1998;   98US-0159812.
PR   15-JAN-1999;   99US-0232149.
PR   09-APR-1999;   99US-0288946.
PR   13-JUL-1999;   99US-0352616.
PR   12-NOV-1999;   99US-0439313.
PR   18-NOV-1999;   99US-0443686.
PR   14-JAN-2000; 2000US-0483672.
PR   27-MAR-2000; 2000US-0536857.
PR   09-MAY-2000; 2000US-0568100.
PR   12-MAY-2000; 2000US-0570737.
PR   13-JUN-2000; 2000US-0593793.
PR   27-JUN-2000; 2000US-0605783.
PR   10-AUG-2000; 2000US-0636215.
PR   29-AUG-2000; 2000US-0651236.
PR   06-SEP-2000; 2000US-0657279.
PR   02-OCT-2000; 2000US-0679426.
PR   10-OCT-2000; 2000US-0685166.
XX
PA   (XUJJ/) XU J.
PA   (DILL/) DILLON D C.
PA   (MITC/) MITCHAM J L.
PA   (HARL/) HARLOCKER S L.
PA   (JIAN/) JIANG Y.
PA   (KALO/) KALOS M D.
PA   (FANG/) FANGER G R.
PA   (RETT/) RETTER M W.
PA   (STOL/) STOLK J A.
PA   (DAYC/) DAY C H.
PA   (VEDV/) VEDVICK T S.
PA   (CART/) CARTER D.
PA   (LISX/) LI S X.
PA   (WANG/) WANG A.
PA   (SKEI/) SKEIKY Y A W.
PA   (HEPL/) HEPLER W T.
PA   (HEND/) HENDERSON R A.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Kalos MD;
PI   Fanger GR,  Retter MW,  Stolk JA,  Day CH,  Vedvick TS,  Carter D;
PI   Li SX,  Wang A,  Skeiky YAW,  Hepler WT,  Henderson RA;
XX
DR   WPI; 2002-255649/30.
XX
PT   New prostate-specific polynucleotides for diagnosing and treating
PT   diseases, in particular prostate cancer, and as markers for the
PT   progression of cancer  -
XX
PS   Claim 1; SEQ ID NO 705; 87pp; English.
XX
CC   The present invention provides prostate-specific coding sequences and
CC   their encoded proteins. These can be used in the diagnosis and treatment
CC   of cancers, particularly prostate cancer. The present sequence is a cDNA
CC   described in the invention.
XX
SQ   Sequence 6976 BP; 1312 A; 1996 C; 1961 G; 1706 T; 1 other;

  Query Match             53.2%;  Score 1815.8;  DB 24;  Length 6976;
  Best Local Similarity   98.7%;  Pred. No. 8.2e-275;
  Matches 1841;  Conservative    0;  Mismatches   22;  Indels    2;  Gaps    1;

Qy       1470 CCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAG 1529
              |||   | ||  ||    ||| |  ||    |    ||||||||||||||||||||||||
Db       5112 CCTGTCCTCTTCCCCTTTTCTTCACCCCTCTGCCTTAGGTGTTCCTGCCCAAATACCGAG 5171

Qy       1530 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTA 1589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5172 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTA 5231

Qy       1590 AGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCAC 1649
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5232 AGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCAC 5291

Qy       1650 CTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTG 1709
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5292 CTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTG 5351

Qy       1710 AGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGG 1769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5352 AGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGG 5411

Qy       1770 ATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGC 1829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5412 ATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGC 5471

Qy       1830 TCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTT 1889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5472 TCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTT 5531

Qy       1890 ACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAA 1949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5532 ACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAA 5591

Qy       1950 CTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTA 2009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5592 CTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTA 5651

Qy       2010 GCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCT 2069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5652 GCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCT 5711

Qy       2070 CTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT 2129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5712 CTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT 5771

Qy       2130 CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTC 2189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5772 CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTC 5831

Qy       2190 TGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT 2249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5832 TGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT 5891

Qy       2250 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG 2309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5892 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG 5951

Qy       2310 TTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGC 2369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5952 TTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGC 6011

Qy       2370 AGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGA 2429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6012 AGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGA 6071

Qy       2430 TTTGAACATATGAC--TTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCA 2487
              |||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db       6072 TTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCA 6131

Qy       2488 GGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCA 2547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6132 GGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCA 6191

Qy       2548 GGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA 2607
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6192 GGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA 6251

Qy        2608 CTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAA 2667
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6252 CTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAA 6311

Qy        2668 TATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGG 2727
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6312 TATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGG 6371

Qy        2728 CTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGG 2787
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6372 CTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGG 6431

Qy        2788 ACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGG 2847
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6432 ACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGG 6491

Qy        2848 TGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCT 2907
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6492 TGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCT 6551

Qy        2908 TCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGG 2967
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6552 TCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGG 6611

Qy        2968 CTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTC 3027
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6612 CTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTC 6671

Qy        3028 CCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 3087
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6672 CCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 6731

Qy        3088 TCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACA 3147
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6732 TCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACA 6791

Qy        3148 GAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTA 3207
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6792 GAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTA 6851

Qy        3208 AGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 3267
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6852 AGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 6911

Qy        3268 GTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAA 3327
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6912 GTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAA 6971

Qy        3328 AAAAA 3332
               |||||
Db        6972 AAAAA 6976
```

```
RESULT 35
ABN81324
ID   ABN81324 standard; cDNA; 1662 BP.
XX
AC   ABN81324;
XX
DT   30-AUG-2002  (first entry)
XX
DE   Human mast cell related splice variant gene MC14 SEQ ID NO 12.
XX
KW   Human; mast cell; MC; antiallergic; antiinflammatory; antiasthmatic;
KW   vasotropic; dermatological; allergy; hypersensitivity; rhinitis; asthma;
KW   gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1662
FT                   /*tag=  a
FT                   /product= "MC14 alternatively spliced variant"
XX
PN   WO200246389-A2.
XX
PD   13-JUN-2002.
XX
PF   07-DEC-2001; 2001WO-US46180.
XX
PR   08-DEC-2000; 2000US-251835P.
PR   14-MAR-2001; 2001US-275479P.
PR   28-MAR-2001; 2001US-279115P.
PR   02-APR-2001; 2001US-280143P.
XX
PA   (UNIO ) UCB SA.
XX
PI   Nocka K,  Pirozzi G,  Einstein R;
XX
DR   WPI; 2002-508560/54.
DR   P-PSDB; ABB77575.
XX
PT   Novel isolated nucleic acids that are differentially expressed in mast
PT   cells in patients with allergic hypersensitivity, encoding proteins
PT   associated with mast cell regranulation and allergic hypersensitivity
PT   -
XX
PS   Claim 1; Page 115-117; 119pp; English.
XX
CC   The invention relates to isolated nucleic acid (ABN81319-ABN81324),
CC   corresponding to genes differentially expressed in mast cells following
CC   activation or in patients with allergic hypersensitivity disease, (I)
CC   that encodes proteins (ABB77569-ABB77575) (II) or a protein fragment of
CC   (II) if at least 6 amino acids. (II) is useful for identifying binding
CC   partners. (I) or (II) is useful for diagnosing or treating a disease
CC   state (e.g. allergic hypersensitivity, seasonal rhinitis, asthma,
CC   urticaria or atopic dermatitis or mastocytosis) in a subject which
CC   involves determining the level of expression of (I) or (II). A computer
CC   system, comprising a database containing information identifying the
CC   expression level in a tissue or at least one mast cell of (I), is useful
CC   for presenting information to identify the relative expression level of
CC   (I). (II) is used as a marker to detect, diagnose or identify an allergic
CC   response in a patient. The protein can also serve as a target that
CC   modulate gene expression or activity and as an antigen to raise
CC   polyclonal or monoclonal antibodies. (II) is useful for identifying
CC   agents that modulate expression of the protein or agents, such as
CC   agonists or antagonists. The agonists or antagonists are useful for
CC   modulating biological activity and function of (II) and thus are useful
CC   for alleviating disease conditions such as allergic hypersensitivity,
CC   seasonal rhinitis, asthma, urticaria, atopic dermatitis or mastocytosis.
XX
SQ   Sequence 1662 BP; 229 A; 551 C; 511 G; 371 T; 0 other;

  Query Match             48.7%;  Score 1662;  DB 24;  Length 1662;
  Best Local Similarity   100.0%;  Pred. No. 7.5e-251;
  Matches 1662;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 60

Qy         344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 120

Qy         404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 180

Qy         464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 240
```

```
Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 300

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 360

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 420

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 480

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 540

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 600

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 660

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 720

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 780

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 840

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 900

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 960

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1020

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1080

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1140

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1200

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1260

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1320

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1380

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1440

Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1500

Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1560

Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1620

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAG 1945
              ||||||||||||||||||||||||||||||||||||||||||
Db       1621 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAG 1662
```

```
RESULT 36
ABA91283
ID   ABA91283 standard; DNA; 2133 BP.
XX
AC   ABA91283;
XX
DT   08-APR-2002  (first entry)
XX
DE   Thioredoxin-ubiquitin-P501S(aa55-553)His triple gene fusion.
XX
KW   Thioredoxin; trxA; ubiquitin; P501S; tumour; prostate; antigen;
KW   cancer; vaccine; therapy; human; gene; ds.
XX
OS   Chimeric - Homo sapiens.
OS   Chimeric - Escherichia coli.
XX
PN   WO200200892-A1.
XX
PD   03-JAN-2002.
XX
PF   19-JUN-2001; 2001WO-EP06952.
XX
PR   26-JUN-2000; 2000GB-0015619.
PR   30-OCT-2000; 2000GB-0026484.
XX
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Cabezon Silva TEV,  Delisse AEF;
XX
DR   WPI; 2002-147888/19.
DR   P-PSDB; AAM50661.
XX
PT   Novel DNA sequence encoding triple fusion protein comprising ubiquitin
PT   fused between thioredoxin and polypeptide of interest, useful for
PT   producing recombinant polypeptide of interest suitable for medicinal
PT   use  -
XX
PS   Example 5; Fig 9a; 87pp; English.
XX
CC   The present sequence is that of a triple gene fusion comprising,
CC   from the 5' end, the trxA thioredoxin gene from Escherichia coli,
CC   the human ubiquitin coding sequence, the prostate antigen P501S
CC   (amino acids 55-553) coding sequence, and DNA encoding a histidine
CC   tail.  The triple fusion was constructed in plasmid pRIT15063,
CC   which included the Saccharomyces cerevisiae CUP1 promoter and
CC   yeast alpha prepro signal sequence.  The triple fusion protein
CC   (see AAM50661) was produced in E. coli GI724 transformants.  This
CC   is an example of the production of triple fusion proteins of the
CC   invention comprising ubiquitin fused between thioredoxin and a
```

```
CC   protein of interest, in this case P501S(aa55-553).  A claimed method
CC   of producing a recombinant protein of interest involves: culturing
CC   a host cell (preferably E. coli) under conditions which allow
CC   for co-expression of the triple fusion and a ubiquitin-specific
CC   endoprotease (especially UBP1 from Saccharomyces cerevisiae);
CC   and recovering the recombinant protein directly from the bacterial
CC   cells after it has been subjected to the action of the
CC   ubiquitin-specific endoprotease in vivo.  In the present case,
CC   a p501-like protein of 509 amino acids is generated.  The
CC   recombinant protein can used as a vaccine for cancer therapy.
XX
SQ   Sequence 2133 BP; 384 A; 656 C; 614 G; 479 T; 0 other;

  Query Match             44.1%;  Score 1503;  DB 24;  Length 2133;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-226;
  Matches 1503;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        440 ATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGC 499
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 ATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGC 663

Qy        500 TCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTG 559
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 TCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTG 723

Qy        560 TCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTG 619
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        724 TCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTG 783

Qy        620 CTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTG 679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        784 CTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTG 843

Qy        680 GACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGG 739
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        844 GACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGG 903

Qy        740 GACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGC 799
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        904 GACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGC 963

Qy        800 TGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTG 859
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        964 TGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTG 1023

Qy        860 GGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCA 919
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1024 GGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCA 1083

Qy        920 GCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTG 979
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1084 GCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTG 1143

Qy        980 TCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAAC 1039
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1144 TCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAAC 1203

Qy       1040 CTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGC 1099
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1204 CTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGC 1263

Qy       1100 CGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTAC 1159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1264 CGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTAC 1323

Qy       1160 ACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAG 1219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1324 ACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAG 1383

Qy       1220 GCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGC 1279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1384 GCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGC 1443

Qy       1280 GCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGA 1339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1444 GCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGA 1503

Qy       1340 GCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCC 1399
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1504 GCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCC 1563

Qy       1400 CACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTG 1459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1564 CACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTG 1623

Qy       1460 CAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCC 1519
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1624 CAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCC 1683

Qy       1520 AAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTG 1579
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1684 AAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTG 1743

Qy       1580 CCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGC 1639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1744 CCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGC 1803

Qy       1640 CTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTG 1699
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1804 CTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTG 1863

Qy       1700 GTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTC 1759
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1864 GTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTC 1923

Qy       1760 GCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCC 1819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1924 GCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCC 1983

Qy       1820 ATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTG 1879
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1984 ATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTG 2043

Qy       1880 GTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCA 1939
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2044 GTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCA 2103

Qy       1940 GCG 1942
              |||
Db       2104 GCG 2106

RESULT 37
ABA91284
ID   ABA91284 standard; DNA; 1593 BP.
XX
AC   ABA91284;
XX
DT   08-APR-2002  (first entry)
XX
DE   Thioredoxin-ubiquitin-P501S(aa1-320)His triple gene fusion.
XX
KW   Thioredoxin; trxA; ubiquitin; P501S; tumour; prostate; antigen;
KW   cancer; vaccine; therapy; human; gene; ds.
XX
OS   Chimeric - Homo sapiens.
OS   Chimeric - Escherichia coli.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1593
FT                   /*tag=  a
FT                   /transl_except= (pos:607..609,aa:Xaa)
FT                   /note= "Xaa not given in amino acid sequence of
FT                          Figure 12b (AAM50662), which has a Val
FT                          residue at amino acid position 208 not
FT                          not encoded by the present sequence"
```

```
XX
PN   WO200200892-A1.
XX
PD   03-JAN-2002.
XX
PF   19-JUN-2001; 2001WO-EP06952.
XX
PR   26-JUN-2000; 2000GB-0015619.
PR   30-OCT-2000; 2000GB-0026484.
XX
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Cabezon Silva TEV,  Delisse AEF;
XX
DR   WPI; 2002-147888/19.
DR   P-PSDB; AAM50662.
XX
PT   Novel DNA sequence encoding triple fusion protein comprising ubiquitin
PT   fused between thioredoxin and polypeptide of interest, useful for
PT   producing recombinant polypeptide of interest suitable for medicinal
PT   use -
XX
PS   Example 8; Fig 12b; 87pp; English.
XX
CC   The present sequence is that of a triple gene fusion comprising,
CC   from the 5' end, the trxA thioredoxin gene from Escherichia coli,
CC   the human ubiquitin coding sequence, the prostate antigen P501S
CC   (amino acids 1-320) coding sequence, and DNA encoding a histidine
CC   tail.  The triple fusion was constructed in plasmid pRIT15115,
CC   under the control of the lambda pL promoter.  Triple fusion
CC   protein (see AAM50662) was produced in E. coli GI724 transformants.
CC   This is an example of the production of triple fusion proteins of
CC   the invention comprising ubiquitin fused between thioredoxin and a
CC   protein of interest, in this case P501S(aa1-320).  A claimed method
CC   of producing a recombinant protein of interest involves: culturing
CC   a host cell (preferably E. coli) under conditions which allow
CC   for co-expression of the triple fusion and a ubiquitin-specific
CC   endoprotease (especially UBP1 from Saccharomyces cerevisiae);
CC   and recovering the recombinant protein directly from the bacterial
CC   cells after it has been subjected to the action of the
CC   ubiquitin-specific endoprotease in vivo.  In the present case,
CC   expression was controlled by addition of tryptophan.  The
CC   recombinant protein can used as a vaccine for cancer therapy.
XX
SQ   Sequence 1593 BP; 308 A; 472 C; 461 G; 352 T; 0 other;

  Query Match             28.2%;  Score 961.2;  DB 24;  Length 1593;
  Best Local Similarity   99.7%;  Pred. No. 2.2e-141;
  Matches  963;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        283 TATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCT 342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 TATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCT 662

Qy        343 GGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCC 402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 GGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCC 722

Qy        403 GCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGG 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 GCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGG 782

Qy        463 TCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGG 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 TCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGG 842

Qy        523 ACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCT 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 ACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCT 902

Qy        583 CTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCT 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        903 CTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCT 962

Qy        643 GGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTT 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 GGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTT 1022

Qy        703 CACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGC 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 CACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGC 1082

Qy        763 CTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1083 CTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGC 1142

Qy        823 CATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1143 CATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTT 1202

Qy        883 TGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGA 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1203 TGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGA 1262

Qy        943 GGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1263 GGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCA 1322

Qy       1003 CTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCT 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1323 CTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCT 1382

Qy       1063 GCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTG 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1383 GCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTG 1442

Qy       1123 CAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1443 CAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCT 1502

Qy       1183 GTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1503 GTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGG 1562

Qy       1243 CGTTCG 1248
              |  | |
Db       1563 CACTAG 1568


RESULT 38
AAD05264/c
ID   AAD05264 standard; cDNA; 1019 BP.
XX
AC   AAD05264;
XX
DT   18-JUL-2001  (first entry)
XX
DE   Human secreted protein-encoding gene 11 cDNA clone HWBAR14, SEQ ID NO:55.
XX
KW   Human; secreted protein; proliferative disorder; cancer; tumour; asthma;
KW   foetal abnormality; developmental abnormality; haematopoietic disorder;
KW   immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW   Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;
KW   psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;
KW   inflammation; neurological disorder; Alzheimer's disease; food additive;
KW   angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;
KW   pregnancy-related disorder; endocrine disorder; infection; wound healing;
KW   cell culture; chemotaxis; vulnerary; binding partner identification;
KW   gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             complement (528..139)
FT                   /*tag=  a
FT                   /product= "Human secreted protein"
```

```
FT                   /note= "Does not include start codon"
FT                   /partial
FT   sig_peptide     complement (528..526)
FT                   /*tag= b
FT   mat_peptide     complement (525..142)
FT                   /*tag= c
FT                   /product= "Mature human secreted protein"
XX
PN   WO200134629-A1.
XX
PD   17-MAY-2001.
XX
PF   08-NOV-2000; 2000WO-US30654.
XX
PR   12-NOV-1999;   99US-0164835.
PR   27-JUL-2000; 2000US-0221142.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Komatsoulis GA,  Wei P,  Fiscella M,  Baker KP;
XX
DR   WPI; 2001-308779/32.
DR   P-PSDB; AAE01396.
XX
PT   New nucleic acid encoding one of 21 human secreted proteins for
PT   diagnosing, preventing, treating or ameliorating medical conditions,
PT   such as autoimmune disease and cancer, and used as a food additive or
PT   preservative -
XX
PS   Claim 1; Page 414-415; 490pp; English.
XX
CC   AAD05220-AAD05282 represent cDNAs corresponding to 21 human secreted
CC   protein genes, and AAE01352-AAE01413 represent the proteins they encode.
CC   AAE01415-AAE01433 represent human secreted protein fragments or variants.
CC   The secreted proteins and their genes are useful for preventing, treating
CC   or ameliorating medical conditions, e.g., by protein or gene therapy.
CC   Pathological conditions can be diagnosed by determining the amount of the
CC   new protein in a sample or by determining the presence of mutations in
CC   the new genes. Specific uses are described for each of the 21 genes,
CC   based on the tissues in which they are most highly expressed, and include
CC   developing products for the diagnosis or treatment of proliferative
CC   disorders, cancer, tumours, foetal and developmental abnormalities,
CC   haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
CC   diseases (e.g., rheumatoid arthritis), inflammation, allergies,
CC   neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
CC   cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
CC   psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
CC   angiogenic disorders, kidney disorders, gastrointestinal disorders,
CC   pregnancy-related disorders, endocrine disorders, and infections. The
CC   proteins can also be used to aid wound healing and epithelial cell
CC   proliferation, to prevent skin aging due to sunburn, to maintain organs
CC   before transplantation, for supporting cell culture of primary tissues,
CC   to regenerate tissues, to identify their cognate ligands or binding
CC   partners, and in chemotaxis, and can be used as a food additive or
CC   preservative to modify storage properties. Antibodies specific for a
CC   protein of the invention can be used in alleviating symptoms associated
CC   with the disorders mentioned above, and in diagnostic immunoassays e.g.,
CC   radioimmunoassay or enzyme linked immunosorbent assay (ELISA).
CC   The present sequence represents a human secreted protein-encoding cDNA of
CC   the invention.
XX
SQ   Sequence 1019 BP; 276 A; 227 C; 257 G; 255 T; 4 other;

  Query Match             27.1%;  Score 925.2;  DB 22;  Length 1019;
  Best Local Similarity   98.3%;  Pred. No. 8.7e-136;
  Matches  941;  Conservative    3;  Mismatches   12;  Indels    1;  Gaps    1;

Qy       2394 ATGGGAGTTTCTAGGATGAAAC-ACTCCTCCATGGGATTTGAACATATGACTTATTTGTA 2452
              |||||||||||||||||||  | |||| |||||||||||||||||||| |  |  ||
Db        989 ATGGGAGTTTCTAGGATGAAGCGACTCNTCCATGGGATTTGAACATATAAAGTTATTTGT 930

Qy       2453 GGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTC 2512
              |||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        929 AGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTSAGCCCACAGCACTGTC 870

Qy       2513 TTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGT 2572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        869 TTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGT 810

Qy       2573 TGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGG 2632
              ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        809 TGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTARCAAAGTAGAAGG 750

Qy       2633 GAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCC 2692
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        749 GAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCC 690

Qy       2693 AACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCT 2752
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 AACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCT 630

Qy       2753 CCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAA 2812
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 CCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAA 570

Qy       2813 ATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGG 2872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 ATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGG 510

Qy       2873 CTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCC 2932
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 CTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCC 450

Qy       2933 CACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTT 2992
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 CACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTT 390

Qy       2993 CCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTG 3052
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 CCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTG 330

Qy       3053 GAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCC 3112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 GAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCC 270

Qy       3113 CCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTG 3172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 CCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTG 210

Qy       3173 AGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCT 3232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 AGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCT 150

Qy       3233 TATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTA 3292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 TATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTA 90

Qy       3293 TGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAA 3349
              ||||||||||||||||||||:||||||||| |||||||||||||||||||||||
Db         89 TGGTGACAAAATTAAAGGCTTTYTTATATGTTAAAAAAAAAAAAAAAAAAAAAAAAA 33


RESULT 39
AAS64153
ID   AAS64153 standard; cDNA; 1203 BP.
XX
AC   AAS64153;
XX
DT   29-JAN-2002  (first entry)
XX
DE   Human /M. tuberculosis fusion protein RA12-P501S-E2 cDNA.
XX
KW   Human; prostate cancer; ss; cytostatic; immunostimulant; tumour.
XX
OS   Chimeric - Homo sapiens.
OS   Chimeric - Microbacterium tuberculosis.
```

```
XX
PN   WO200173032-A2.
XX
PD   04-OCT-2001.
XX
PF   27-MAR-2001; 2001WO-US09919.
XX
PR   27-MAR-2000; 2000US-0536857.
PR   09-MAY-2000; 2000US-0568100.
PR   12-MAY-2000; 2000US-0570737.
PR   13-JUN-2000; 2000US-0593793.
PR   27-JUN-2000; 2000US-0605783.
PR   10-AUG-2000; 2000US-0636215.
PR   29-AUG-2000; 2000US-0651236.
PR   06-SEP-2000; 2000US-0657279.
PR   02-OCT-2000; 2000US-0679426.
PR   10-OCT-2000; 2000US-0685166.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Kalos MD;
PI   Fanger GR,  Retter MW,  Stolk JA,  Day CH,  Vedvick TS,  Carter D;
PI   Li SX,  Wang A,  Skeiky YAW,  Hepler WT,  Henderson RA;
XX
DR   WPI; 2001-639232/73.
DR   P-PSDB; AAU69907.
XX
PT   New human prostate-specific polypeptides and polynucleotides useful for
PT   the diagnosis and treatment of cancer, especially prostate cancer -
XX
PS   Example 17; Page 542-543; 579pp; English.
XX
CC   The invention relates to isolated prostate-specific
CC   polynucleotides, polypeptides, fusion proteins of the polypeptides,
CC   antibodies raised against the polypeptides (or antigenic epitopes
CC   derived from them) and antigen-presenting cells expressing the
CC   polypeptides. The antibodies are useful for detecting the presence of
CC   cancer, especially prostate cancer. The polypeptides, polynucleotides and
CC   the antigen-presenting cells are useful for stimulating and/or expanding
CC   T cells specific for a tumour protein, and for inhibiting the development
CC   of cancer especially prostate cancer. Compositions comprising the
CC   polynucleotide and/or polypeptide are useful for stimulating an immune
CC   response, and for treating cancer. The oligonucleotide is useful for
CC   detecting cancer. The present sequence is a prostate specific
CC   polynucleotide of the invention.
XX
SQ   Sequence 1203 BP; 175 A; 415 C; 368 G; 245 T; 0 other;

  Query Match            23.3%;  Score 794.6;  DB 22;  Length 1203;
  Best Local Similarity  99.5%;  Pred. No. 2.2e-115;
  Matches  797;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        379 GGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTT 438
              ||||| |  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 GGCCGAATTCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTT 461

Qy        439 CATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGG 498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 CATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGG 521

Qy        499 CTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACT 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 CTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACT 581

Qy        559 GTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCT 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 GTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCT 641

Qy        619 GCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCT 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        642 GCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCT 701

Qy        679 GGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCG 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        702 GGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCG 761

Qy        739 GGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGG 798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        762 GGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGG 821

Qy        799 CTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCT 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        822 CTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCT 881

Qy        859 GGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGC 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        882 GGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGC 941

Qy        919 AGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCT 978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        942 AGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCT 1001

Qy        979 GTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAA 1038
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1002 GTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAA 1061

Qy       1039 CCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCG 1098
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1062 CCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCG 1121

Qy       1099 CCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTA 1158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1122 CCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTA 1181

Qy       1159 CACGGATTTCGTGGGCGAGGG 1179
              ||||||||||||||||||| |
Db       1182 CACGGATTTCGTGGGCGAGTG 1202


RESULT 40
AAH93917
ID   AAH93917 standard; cDNA; 1203 BP.
XX
AC   AAH93917;
XX
DT   04-OCT-2001  (first entry)
XX
DE   Ra12-P501S-E2 construct cDNA sequence.
XX
KW   Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW   cytostatic; gene therapy; metastasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200151633-A2.
XX
PD   19-JUL-2001.
XX
PF   16-JAN-2001; 2001WO-US01574.
XX
PR   14-JAN-2000; 2000US-0483672.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Reed SG;
PI   Kalos MD,  Fanger GR,  Day CH,  Retter MW,  Stolk JA,  Skeiky YAW;
PI   Wang A,  Meagher MJ;
XX
DR   WPI; 2001-425873/45.
XX
PT   New polynucleotide encoding a prostate-specific protein, for
PT   diagnosing, monitoring and treating prostate cancer in a patient and
PT   for use in vaccines -
XX
PS   Claim 8; Page 504; 543pp; English.
XX
```

```
CC   The present invention describes polynucleotide sequences (I) which encode
CC   prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC   and can be used in vaccine production and gene therapy. (I), (II),
CC   antibodies to (II), fusion proteins comprising (II), and isolated
CC   T cells prepared using (I) or (II) are used treat cancer in a patient.
CC   (I) and the antibodies are also used in the detection of cancer in a
CC   patient. The cancer that is diagnosed or treated is particularly
CC   prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
CC   (I) can be used for monitoring the progression of cancer in a patient.
CC   (I) and (II) can also be used to improve diagnostic and therapeutic
CC   methods for prostate cancer. They can indicate the level of metastasis
CC   as well as the prostate volume. AAH93357 to AAH93944 and AAM01115 to
CC   AAM01318 represent polynucleotide and amino acid sequences used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 1203 BP; 175 A; 415 C; 368 G; 245 T; 0 other;

  Query Match             23.3%;  Score 794.6;  DB 22;  Length 1203;
  Best Local Similarity   99.5%;  Pred. No. 2.2e-115;
  Matches  797;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        379 GGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTT 438
              ||||| |  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 GGCCGAATTCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTT 461

Qy        439 CATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGG 498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 CATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGG 521

Qy        499 CTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACT 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 CTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACT 581

Qy        559 GTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCT 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 GTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCT 641

Qy        619 GCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCT 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        642 GCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCT 701

Qy        679 GGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCG 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        702 GGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCG 761

Qy        739 GGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGG 798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        762 GGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGG 821

Qy        799 CTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCT 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        822 CTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCT 881

Qy        859 GGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGC 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        882 GGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGC 941

Qy        919 AGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCT 978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        942 AGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCT 1001

Qy        979 GTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAA 1038
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1002 GTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAA 1061

Qy       1039 CCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCG 1098
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1062 CCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCG 1121

Qy       1099 CCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTA 1158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1122 CCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTA 1181

Qy       1159 CACGGATTTCGTGGGCGAGGG 1179
              ||||||||||||||||||| |
Db       1182 CACGGATTTCGTGGGCGAGTG 1202


RESULT 41
ABL95524
ID   ABL95524 standard; cDNA; 1203 BP.
XX
AC   ABL95524;
XX
DT   19-JUL-2002  (first entry)
XX
DE   Ra12-P501S-E2 construct cDNA sequence SEQ ID NO 851.
XX
KW   Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
KW   gene therapy; gene; ss.
XX
OS   Chimeric - Mycobacterium tuberculosis.
OS   Chimeric - Homo sapiens.
XX
PN   US2002022248-A1.
XX
PD   21-FEB-2002.
XX
PF   12-JAN-2001; 2001US-0759143.
XX
PR   25-FEB-1997;   97US-0806099.
PR   01-AUG-1997;   97US-0904804.
PR   09-FEB-1998;   98US-0020956.
PR   25-FEB-1998;   98US-0030607.
PR   14-JUL-1998;   98US-0115453.
PR   23-SEP-1998;   98US-0159812.
PR   15-JAN-1999;   99US-0232149.
PR   09-APR-1999;   99US-0288946.
PR   13-JUL-1999;   99US-0352616.
PR   12-NOV-1999;   99US-0439313.
PR   18-NOV-1999;   99US-0443686.
PR   14-JAN-2000; 2000US-0483672.
PR   27-MAR-2000; 2000US-0536857.
PR   09-MAY-2000; 2000US-0568100.
PR   12-MAY-2000; 2000US-0570737.
PR   13-JUN-2000; 2000US-0593793.
PR   27-JUN-2000; 2000US-0605783.
PR   10-AUG-2000; 2000US-0636215.
PR   29-AUG-2000; 2000US-0651236.
PR   06-SEP-2000; 2000US-0657279.
PR   02-OCT-2000; 2000US-0679426.
PR   10-OCT-2000; 2000US-0685166.
XX
PA   (XUJJ/) XU J.
PA   (DILL/) DILLON D C.
PA   (MITC/) MITCHAM J L.
PA   (HARL/) HARLOCKER S L.
PA   (JIAN/) JIANG Y.
PA   (KALO/) KALOS M D.
PA   (FANG/) FANGER G R.
PA   (RETT/) RETTER M W.
PA   (STOL/) STOLK J A.
PA   (DAYC/) DAY C H.
PA   (VEDV/) VEDVICK T S.
PA   (CART/) CARTER D.
PA   (LISX/) LI S X.
PA   (WANG/) WANG A.
PA   (SKEI/) SKEIKY Y A W.
PA   (HEPL/) HEPLER W T.
PA   (HEND/) HENDERSON R A.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Kalos MD;
PI   Fanger GR,  Retter MW,  Stolk JA,  Day CH,  Vedvick TS,  Carter D;
PI   Li SX,  Wang A,  Skeiky YAW,  Hepler WT,  Henderson RA;
XX
DR   WPI; 2002-255649/30.
```

```
XX
PT   New prostate-specific polynucleotides for diagnosing and treating
PT   diseases, in particular prostate cancer, and as markers for the
PT   progression of cancer -
XX
PS   Example 17; SEQ ID NO 851; 87pp; English.
XX
CC   The present invention provides prostate-specific coding sequences and
CC   their encoded proteins. These can be used in the diagnosis and treatment
CC   of cancers, particularly prostate cancer. The present sequence is a cDNA
CC   described in the invention.
XX
SQ   Sequence 1203 BP; 175 A; 415 C; 368 G; 245 T; 0 other;

  Query Match             23.3%;  Score 794.6;  DB 24;  Length 1203;
  Best Local Similarity   99.5%;  Pred. No. 2.2e-115;
  Matches  797;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        379 GGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGTAGAGGAGAAGTT 438
              ||||| |  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 GGCCGAATTCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGTAGAGGAGAAGTT 461

Qy        439 CATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGG 498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 CATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGG 521

Qy        499 CTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACT 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 CTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACT 581

Qy        559 GTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCT 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 GTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCT 641

Qy        619 GCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCT 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        642 GCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCT 701

Qy        679 GGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCG 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        702 GGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCG 761

Qy        739 GGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGG 798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        762 GGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGG 821

Qy        799 CTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCT 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        822 CTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCT 881

Qy        859 GGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGC 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        882 GGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGC 941

Qy        919 AGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCT 978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        942 AGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCT 1001

Qy        979 GTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAA 1038
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1002 GTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAA 1061

Qy       1039 CCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCG 1098
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1062 CCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCG 1121

Qy       1099 CCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTA 1158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1122 CCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTA 1181

Qy       1159 CACGGATTTCGTGGGCGAGGG 1179
              ||||||||||||||||||| |
Db       1182 CACGGATTTCGTGGGCGAGTG 1202
```

```
RESULT 42
AAV61144
ID   AAV61144 standard; cDNA; 789 BP.
XX
AC   AAV61144;
XX
DT   06-JAN-1999  (first entry)
XX
DE   3' cDNA sequence of prostate tumour clone L1-12.
XX
KW   Prostate; cancer; tumour; vaccine; immunogen; clone; ss.
XX
OS   Homo sapiens.
XX
PN   WO9837093-A2.
XX
PD   27-AUG-1998.
XX
PF   25-FEB-1998;   98WO-US03492.
XX
PR   09-FEB-1998;   98US-0020956.
PR   25-FEB-1997;   97US-0806099.
PR   01-AUG-1997;   97US-0904804.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Dillon DC,  Xu J;
XX
DR   WPI; 1998-609886/51.
XX
PT   Polypeptides comprising immunogenic portions of prostate proteins -
PT   used in a vaccine for the treatment of prostate cancer
XX
PS   Claim 3; Page 38-39; 130pp; English.
XX
CC   The present sequence is a new DNA which encodes an immunogenic portion
CC   of a prostate tumour protein. The encoded immunogen, or the DNA itself,
CC   can be used as a vaccine for the treatment of prostate cancer. The DNA
CC   was identified by analysis of a subtracted cDNA library obtained by
CC   subtracting a prostate tumour cDNA expression library with a normal
CC   tissue cDNA library.
XX
SQ   Sequence 789 BP; 122 A; 250 C; 221 G; 182 T; 14 other;

  Query Match             19.7%;  Score 673.4;  DB 19;  Length 789;
  Best Local Similarity   94.2%;  Pred. No. 1.8e-96;
  Matches  745;  Conservative    0;  Mismatches   40;  Indels    6;  Gaps    5;

Qy       1341 CAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCC 1400
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGTCTATNTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCC 60

Qy       1401 ACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGC 1460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGC 120

Qy       1461 AGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCA 1520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCA 180

Qy       1521 AATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC 1580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC 240

Qy       1581 CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCC 1640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCC 300

Qy       1641 TGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGG 1700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGG 360
```

```
Qy       1701 TGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCG 1760
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        361 TGGTGGGTGAGCCCACCGANGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCG 420

Qy       1761 CCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCA 1820
              ||||||||||||||| ||||||||||||||| ||||||||||||||||||||||||||||
Db        421 CCATCCTGGATAGTG-CTTCCTGCTGTCCCANGTGGCCCCATCCCTGTTTATGGGCTCCA 479

Qy       1821 TTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGG 1880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 TTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGG 539

Qy       1881 TCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAG 1940
              || ||||||||||||||||||||||| |||||||||||| ||| ||||||||||||||||
Db        540 TC-CCATTTACTTTGCTACACAGGTANTATTTGACAAGAACGANTTGGCCAAATACTCAG 598

Qy       1941 CGTAGAAAACTTCCAGCA--CATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCC 1998
              |||  |||| ||||||||      | |||||||| |||||||||||||||| ||| |||||
Db        599 CGTTAAAAAATTCCAGCAACATTGGGGGTGGAAGGCCTGCCTCACTGGGT-CCAACTCCC 657

Qy       1999 CGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGT 2058
              ||||||||||| |||||||||||||||||  ||||||||| ||||||||||||||||| |
Db        658 CGCTCCTGTTAACCCCATGGGGCTGCCGGCTTGGCCGCCAATTTCTGTTGCTGCCAAANT 717

Qy       2059 AATGTGGCTCTCTGCTGCCACCCTGTGCT-GCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2117
               |||||||||||||||||||    |||| ||||| |||| ||          |||||
Db        718 NATGTGGCTCTCTGCTGCCACCTGTTGCTGGCTGAAGTGCNTACNGCNCANCTNGGGGGG 777

Qy       2118 TGGGGCGTCCC 2128
              | ||| || ||
Db        778 TNGGGNGTTCC 788
```

```
RESULT 43
AAV58487
ID   AAV58487 standard; cDNA; 789 BP.
XX
AC   AAV58487;
XX
DT   08-DEC-1998  (first entry)
XX
DE   3' fragment of prostate tumour specific gene L1-12.
XX
KW   Prostate tumour specific gene; human; prostate cancer; detection;
KW   therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO9837418-A2.
XX
PD   27-AUG-1998.
XX
PF   25-FEB-1998;   98WO-US03690.
XX
PR   09-FEB-1998;   98US-0904809.
PR   25-FEB-1997;   97US-0806596.
PR   01-AUG-1997;   97US-0904809.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Dillon DC,  Xu J;
XX
DR   WPI; 1998-480805/41.
XX
PT   Novel human prostate specific tumour protein and fragments - useful
PT   for detecting and treating prostate cancers
XX
PS   Claim 1; Page 39; 141pp; English.
XX
CC   This sequence represents a human prostate tumour specific gene, and can
CC   be used in the method of the invention. The method is for detecting
CC   prostate cancer comprises contacting a biological sample with an agent
CC   able to bind an immunogenic portion of a prostate protein (such as
CC   encoded by this sequence). An antibody which binds to an immunogenic
CC   portion of the prostate protein, and the method can be used to detect,
CC   monitor progression of, or treat prostate cancers. The antibody may
CC   also be conjugated to a therapeutic agent for use in therapy of prostate
CC   cancers.
XX
SQ   Sequence 789 BP; 122 A; 250 C; 221 G; 182 T; 14 other;

  Query Match             19.7%;  Score 673.4;  DB 19;  Length 789;
  Best Local Similarity   94.2%;  Pred. No. 1.8e-96;
  Matches  745;  Conservative    0;  Mismatches   40;  Indels    6;  Gaps    5;

Qy       1341 CAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCC 1400
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGTCTATNTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCC 60

Qy       1401 ACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGC 1460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGC 120

Qy       1461 AGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCA 1520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCA 180

Qy       1521 AATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC 1580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC 240

Qy       1581 CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCC 1640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCC 300

Qy       1641 TGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGG 1700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGG 360

Qy       1701 TGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCG 1760
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        361 TGGTGGGTGAGCCCACCGANGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCG 420

Qy       1761 CCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCA 1820
              ||||||||||||||| ||||||||||||||| ||||||||||||||||||||||||||||
Db        421 CCATCCTGGATAGTG-CTTCCTGCTGTCCCANGTGGCCCCATCCCTGTTTATGGGCTCCA 479

Qy       1821 TTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGG 1880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 TTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGG 539

Qy       1881 TCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAG 1940
              || ||||||||||||||||||||||| |||||||||||| ||| ||||||||||||||||
Db        540 TC-CCATTTACTTTGCTACACAGGTANTATTTGACAAGAACGANTTGGCCAAATACTCAG 598

Qy       1941 CGTAGAAAACTTCCAGCA--CATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCC 1998
              |||  |||| ||||||||      | |||||||| |||||||||||||||| ||| |||||
Db        599 CGTTAAAAAATTCCAGCAACATTGGGGGTGGAAGGCCTGCCTCACTGGGT-CCAACTCCC 657

Qy       1999 CGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGT 2058
              ||||||||||| |||||||||||||||||  ||||||||| ||||||||||||||||| |
Db        658 CGCTCCTGTTAACCCCATGGGGCTGCCGGCTTGGCCGCCAATTTCTGTTGCTGCCAAANT 717

Qy       2059 AATGTGGCTCTCTGCTGCCACCCTGTGCT-GCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2117
               |||||||||||||||||||    |||| ||||| |||| ||          |||||
Db        718 NATGTGGCTCTCTGCTGCCACCTGTTGCTGGCTGAAGTGCNTACNGCNCANCTNGGGGGG 777

Qy       2118 TGGGGCGTCCC 2128
              | ||| || ||
Db        778 TNGGGNGTTCC 788

RESULT 44
AAA06250
ID   AAA06250 standard; cDNA; 789 BP.
```

```
XX
AC   AAA06250;
XX
DT   13-JUN-2000  (first entry)
XX
DE   Human immunogenic prostate tumour protein cDNA sequence SEQ ID NO:10.
XX
KW   Human; prostate cancer; diagnosis; tumour; gene therapy; detection;
KW   immunogenic; cytostatic; vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200004149-A2.
XX
PD   27-JAN-2000.
XX
PF   14-JUL-1999;   99WO-US15838.
XX
PR   14-JUL-1998;   98US-0115453.
PR   14-JUL-1998;   98US-0116134.
PR   23-SEP-1998;   98US-0159812.
PR   23-SEP-1998;   98US-0159822.
PR   15-JAN-1999;   99US-0232149.
PR   15-JAN-1999;   99US-0232880.
PR   09-APR-1999;   99US-0288946.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Dillon DC,  Harlocker SL,  Yuqiu J,  Xu J,  Mitcham JL;
XX
DR   WPI; 2000-171268/15.
XX
PT   New polypeptide useful for treating and diagnosing prostate cancer
PT   comprises an immunogenic portion of prostate tumor protein -
XX
PS   Claim 1; Page 99; 263pp; English.
XX
CC   The present invention describes isolated polypeptides, comprising an
CC   immunogenic portion of a prostate tumour protein (PTP). The polypeptides
CC   and polynucleotides encoding them have cytostatic activity and can be
CC   used in vaccines and in gene therapy. The polypeptides and
CC   polynucleotides encoding them, antigen presenting cells which express
CC   the polypeptides, antibodies against the polypeptides and vaccines
CC   comprising them can be used for inhibiting the development of prostate
CC   cancer in a patient. The polypeptides can be used to generate antibodies
CC   or anti-idiotypic antibodies for passive immuno therapy. A portion of
CC   the polynucleotides encoding the polypeptides can be used as a probe or
CC   to modulate the expression of the polypeptides. AAA06241 to AAA06691 and
CC   AAY82000 to AAY82020 represent sequences used in the exemplification of
CC   the present invention.
XX
SQ   Sequence 789 BP; 122 A; 250 C; 221 G; 182 T; 14 other;

  Query Match             19.7%;  Score 673.4;  DB 21;  Length 789;
  Best Local Similarity   94.2%;  Pred. No. 1.8e-96;
  Matches  745;  Conservative    0;  Mismatches   40;  Indels    6;  Gaps    5;

Qy       1341 CAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCC 1400
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGTCTATNTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCC 60

Qy       1401 ACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGC 1460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGC 120

Qy       1461 AGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCA 1520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCA 180

Qy       1521 AATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC 1580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC 240

Qy       1581 CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCC 1640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCC 300

Qy       1641 TGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGG 1700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGG 360

Qy       1701 TGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCG 1760
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        361 TGGTGGGTGAGCCCACCGANGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCG 420

Qy       1761 CCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCA 1820
              ||||||||||||||| ||||||||||||||| ||||||||||||||||||||||||||||
Db        421 CCATCCTGGATAGTG-CTTCCTGCTGTCCCANGTGGCCCCATCCCTGTTTATGGGCTCCA 479

Qy       1821 TTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGG 1880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 TTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGG 539

Qy       1881 TCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAG 1940
              || ||||||||||||||||||||||| |||||||||||| ||| ||||||||||||||||
Db        540 TC-CCATTTACTTTGCTACACAGGTANTATTTGACAAGAACGANTTGGCCAAATACTCAG 598

Qy       1941 CGTAGAAAACTTCCAGCA--CATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCC 1998
              |||  ||||  ||||||||       |  ||||||| ||||||||||||||||| |||  |||||
Db        599 CGTTAAAAAATTCCAGCAACATTGGGGGTGGAAGGCCTGCCTCACTGGGT-CCAACTCCC 657

Qy       1999 CGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGT 2058
              ||||||||||| |||||||||||||||||  ||||||||| ||||||||||||||||| |
Db        658 CGCTCCTGTTAACCCCATGGGGCTGCCGGCTTGGCCGCCAATTTCTGTTGCTGCCAAANT 717

Qy       2059 AATGTGGCTCTCTGCTGCCACCCTGTGCT-GCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2117
               |||||||||||||||||||||   |||| ||||| |||| ||           ||||| 
Db        718 NATGTGGCTCTCTGCTGCCACCTGTTGCTGGCTGAAGTGCNTACNGCNCANCTNGGGGGG 777

Qy       2118 TGGGGCGTCCC 2128
              | ||| || ||
Db        778 TNGGGNGTTCC 788


RESULT 45
AAS63458
ID   AAS63458 standard; cDNA; 789 BP.
XX
AC   AAS63458;
XX
DT   29-JAN-2002  (first entry)
XX
DE   Human prostate cDNA sequence #10.
XX
KW   Human; prostate cancer; ss; cytostatic; immunostimulant; tumour.
XX
OS   Homo sapiens.
XX
PN   WO200173032-A2.
XX
PD   04-OCT-2001.
XX
PF   27-MAR-2001; 2001WO-US09919.
XX
PR   27-MAR-2000; 2000US-0536857.
PR   09-MAY-2000; 2000US-0568100.
PR   12-MAY-2000; 2000US-0570737.
PR   13-JUN-2000; 2000US-0593793.
PR   27-JUN-2000; 2000US-0605783.
PR   10-AUG-2000; 2000US-0636215.
PR   29-AUG-2000; 2000US-0651236.
PR   06-SEP-2000; 2000US-0657279.
PR   02-OCT-2000; 2000US-0679426.
PR   10-OCT-2000; 2000US-0685166.
XX
PA   (CORI-) CORIXA CORP.
```

```
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Kalos MD;
PI   Fanger GR,  Retter MW,  Stolk JA,  Day CH,  Vedvick TS,  Carter D;
PI   Li SX,  Wang A,  Skeiky YAW,  Hepler WT,  Henderson RA;
XX
DR   WPI; 2001-639232/73.
XX
PT   New human prostate-specific polypeptides and polynucleotides useful for
PT   the diagnosis and treatment of cancer, especially prostate cancer -
XX
PS   Claim 1; Page 232-233; 579pp; English.
XX
CC   The invention relates to isolated prostate-specific
CC   polynucleotides, polypeptides, fusion proteins of the polypeptides,
CC   antibodies raised against the polypeptides (or antigenic epitopes
CC   derived from them) and antigen-presenting cells expressing the
CC   polypeptides. The antibodies are useful for detecting the presence of
CC   cancer, especially prostate cancer. The polypeptides, polynucleotides and
CC   the antigen-presenting cells are useful for stimulating and/or expanding
CC   T cells specific for a tumour protein, and for inhibiting the development
CC   of cancer especially prostate cancer. Compositions comprising the
CC   polynucleotide and/or polypeptide are useful for stimulating an immune
CC   response, and for treating cancer. The oligonucleotide is useful for
CC   detecting cancer. The present sequence is a prostate specific
CC   polynucleotide of the invention.
XX
SQ   Sequence 789 BP; 122 A; 250 C; 221 G; 182 T; 14 other;

  Query Match             19.7%;  Score 673.4;  DB 22;  Length 789;
  Best Local Similarity   94.2%;  Pred. No. 1.8e-96;
  Matches  745;  Conservative    0;  Mismatches   40;  Indels    6;  Gaps    5;

Qy       1341 CAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCC 1400
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGTCTATNTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCC 60

Qy       1401 ACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGC 1460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGC 120

Qy       1461 AGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCA 1520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCA 180

Qy       1521 AATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC 1580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC 240

Qy       1581 CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCC 1640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCC 300

Qy       1641 TGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGG 1700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGG 360

Qy       1701 TGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCG 1760
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        361 TGGTGGGTGAGCCCACCGANGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCG 420

Qy       1761 CCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCA 1820
              ||||||||||||||| ||||||||||||||| ||||||||||||||||||||||||||||
Db        421 CCATCCTGGATAGTG-CTTCCTGCTGTCCCANGTGGCCCCATCCCTGTTTATGGGCTCCA 479

Qy       1821 TTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGG 1880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 TTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGG 539

Qy       1881 TCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAG 1940
              || ||||||||||||||||||||||| |||||||||||| ||| ||||||||||||||||
Db        540 TC-CCATTTACTTTGCTACACAGGTANTATTTGACAAGAACGANTTGGCCAAATACTCAG 598

Qy       1941 CGTAGAAAACTTCCAGCA--CATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCC 1998
              |||  ||||  |||||||||      | |||||||| |||||||||||||||| ||| |||||
Db        599 CGTTAAAAAATTCCAGCAACATTGGGGGTGGAAGGCCTGCCTCACTGGGT-CCAACTCCC 657

Qy       1999 CGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGT 2058
              ||||||||||| |||||||||||||||||  ||||||||| ||||||||||||||||| |
Db        658 CGCTCCTGTTAACCCCATGGGGCTGCCGGCTTGGCCGCCAATTTCTGTTGCTGCCAAANT 717

Qy       2059 AATGTGGCTCTCTGCTGCCACCCTGTGCT-GCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2117
                |||||||||||||||||||    |||| ||||| |||| ||            |||||
Db        718 NATGTGGCTCTCTGCTGCCACCTGTTGCTGGCTGAAGTGCNTACNGCNCANCTNGGGGGG 777

Qy       2118 TGGGGCGTCCC 2128
              | ||| || ||
Db        778 TNGGGNGTTCC 788

Search completed: June  9, 2003, 16:13:19
Job time : 816 secs
```

OM nucleic - nucleic search, using sw model

```
Run on:          June  9, 2003, 13:11:18 ; Search time 8483 Seconds
                                           (without alignments)
                                           11698.758 Million cell updates/sec

Title:           US-09-759-143-110
Perfect score:   3410
Sequence:        1 gggaaccagcctgcacgcgc..........aaaaaaataaaaaaaaaaaa 3410

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters:      4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       GenEmbl:*
                 1:  gb_ba:*
                 2:  gb_htg:*
                 3:  gb_in:*
                 4:  gb_om:*
                 5:  gb_ov:*
                 6:  gb_pat:*
                 7:  gb_ph:*
                 8:  gb_pl:*
                 9:  gb_pr:*
                 10:  gb_ro:*
                 11:  gb_sts:*
                 12:  gb_sy:*
                 13:  gb_un:*
                 14:  gb_vi:*
                 15:  em_ba:*
                 16:  em_fun:*
                 17:  em_hum:*
                 18:  em_in:*
                 19:  em_mu:*
                 20:  em_om:*
                 21:  em_or:*
                 22:  em_ov:*
                 23:  em_pat:*
                 24:  em_ph:*
                 25:  em_pl:*
                 26:  em_ro:*
                 27:  em_sts:*
                 28:  em_un:*
                 29:  em_vi:*
                 30:  em_htg_hum:*
                 31:  em_htg_inv:*
                 32:  em_htg_other:*
                 33:  em_htg_mus:*
                 34:  em_htg_pln:*
                 35:  em_htg_rod:*
                 36:  em_htg_mam:*
                 37:  em_htg_vrt:*
                 38:  em_sy:*
                 39:  em_htgo_hum:*
                 40:  em_htgo_mus:*
                 41:  em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3409.6	100.0	3410	6	AX106329	AX106329 Sequence
	2	3409.6	100.0	3410	6	AX140620	AX140620 Sequence
	3	3409.6	100.0	3410	6	AX200480	AX200480 Sequence
	4	3409.6	100.0	3410	6	AX267136	AX267136 Sequence
	5	3409.6	100.0	3410	6	AX429961	AX429961 Sequence
	6	3409.6	100.0	3410	9	AY033593	AY033593 Homo sapi
	7	3292.4	96.6	3320	6	AX327336	AX327336 Sequence
	8	2900.2	85.0	3514	9	AB060851	AB060851 Macaca fa
	9	2585.4	75.8	4034	6	AX200995	AX200995 Sequence
	10	2585.4	75.8	4034	6	AX267730	AX267730 Sequence
	11	2545.6	74.7	2917	9	AB062977	AB062977 Macaca fa
	12	2196.4	64.4	2904	6	AX200994	AX200994 Sequence
	13	2196.4	64.4	2904	6	AX267729	AX267729 Sequence
	14	2142.8	62.8	4894	6	AX200993	AX200993 Sequence
	15	2142.8	62.8	4894	6	AX267728	AX267728 Sequence
	16	2136.4	62.7	2152	6	AR112295	AR112295 Sequence
	17	2114.8	62.0	2143	6	AR112294	AR112294 Sequence
	18	2103.8	61.7	2477	9	HSM804244	AL832933 Homo sapi
	19	1815.8	53.2	6976	6	AX200996	AX200996 Sequence
	20	1815.8	53.2	6976	6	AX267731	AX267731 Sequence
	21	1809.8	53.1	157988	2	AC096533	AC096533 Homo sapi
	22	1583.4	46.4	3354	10	BC031381	BC031381 Mus muscu
	23	1503	44.1	2133	6	AX343857	AX343857 Sequence
	24	1246.2	36.5	2611	10	BC034084	BC034084 Mus muscu
	25	961.2	28.2	1593	6	AX343860	AX343860 Sequence
	26	794.6	23.3	1203	6	AX201078	AX201078 Sequence
	27	794.6	23.3	1203	6	AX267877	AX267877 Sequence
	28	673.4	19.7	789	6	AX106229	AX106229 Sequence
	29	673.4	19.7	789	6	AX140520	AX140520 Sequence
	30	673.4	19.7	789	6	AX200380	AX200380 Sequence
	31	673.4	19.7	789	6	AX267036	AX267036 Sequence
	32	633.2	18.6	198037	2	AC126523	AC126523 Rattus no
c	33	623.6	18.3	178567	2	AC107837	AC107837 Mus muscu
c	34	603	17.7	772	6	AX106230	AX106230 Sequence
c	35	603	17.7	772	6	AX140521	AX140521 Sequence
c	36	603	17.7	772	6	AX200381	AX200381 Sequence
c	37	603	17.7	772	6	AX267037	AX267037 Sequence
	38	442.6	13.0	1593	10	BC024519	BC024519 Mus muscu
c	39	361.8	10.6	198037	2	AC126523	AC126523 Rattus no
	40	319.8	9.4	342	6	AR112287	AR112287 Sequence
	41	291	8.5	294	6	AR112291	AR112291 Sequence
	42	287	8.4	288	6	AR112289	AR112289 Sequence
	43	270	7.9	272	6	AR112290	AR112290 Sequence
	44	265	7.8	265	6	AR112288	AR112288 Sequence
c	45	261.4	7.7	406	11	HSPE54C06	AL033941 H.sapiens

ALIGNMENTS

```
RESULT 1
AX106329
LOCUS       AX106329                3410 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 110 from Patent WO0125272.
ACCESSION   AX106329
VERSION     AX106329.1  GI:13922014
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3410)
  AUTHORS   Xu,J., Skeiky,Y.A., Reed,S.G. and Cheever,M.A.
  TITLE     Compositions and methods for therapy and diagnosis of prostate
            cancer
```

```
JOURNAL    Patent: WO 0125272-A 110 12-APR-2001;
           CORIXA CORPORATION (US)
FEATURES            Location/Qualifiers
     source         1..3410
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
BASE COUNT      667 a   1014 c    945 g    783 t      1 others
ORIGIN

  Query Match             100.0%;  Score 3409.6;  DB 6;  Length 3410;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy        2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy        2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy        2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy        2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy        2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy        2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy        2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy        2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy        2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy        2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy        2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCA 2700
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCA 2700

Qy        2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy        2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy        2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy        2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy        2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy        3001 CCCAACTTTCCCCTACCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3001 CCCAACTTTCCCCTACCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy        3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy        3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy        3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy        3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy        3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy        3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410


RESULT 2
AX140620
LOCUS       AX140620                3410 bp    DNA     linear   PAT 31-MAY-2001
DEFINITION  Sequence 110 from Patent WO0134802.
ACCESSION   AX140620
VERSION     AX140620.1  GI:14280737
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3410)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Retter,M.W., Stolk,J.A., Day,C.H.,
            Skeiky,Y.A. and Wang,A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0134802-A 110 17-MAY-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..3410
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      667 a   1014 c    945 g    783 t      1 others
ORIGIN

  Query Match             100.0%;  Score 3409.6;  DB 6;  Length 3410;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy          61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy         121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy         181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy         241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
```

```
Qy        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy       2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy       2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy       2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy       2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy        2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy        2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy        2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700

Qy        2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy        2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy        2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy        2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy        2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy        3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy        3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy        3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy        3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy        3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy        3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy        3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
```

```
RESULT 3
AX200480
LOCUS       AX200480                3410 bp    DNA     linear   PAT 29-AUG-2001
DEFINITION  Sequence 110 from Patent WO0151633.
ACCESSION   AX200480
VERSION     AX200480.1  GI:15390293
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3410)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Fanger,G.R., Day,C.H., Retter,M.W.,
            Stolk,J.A., Skeiky,Y.A., Wang,A. and Meagher,M.J.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0151633-A 110 19-JUL-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..3410
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      667 a   1014 c    945 g    783 t      1 others
ORIGIN

  Query Match             100.0%;  Score 3409.6;  DB 6;  Length 3410;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy           1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy          61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy         121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy         181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy         241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy         301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy         361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy         421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy         481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy         541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy         601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy         661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy         721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy         781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy       2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy       2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy       2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy       2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700

Qy       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
```

```
Qy       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
```

```
RESULT 4
AX267136
LOCUS       AX267136                3410 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 110 from Patent WO0173032.
ACCESSION   AX267136
VERSION     AX267136.1  GI:16515920
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
            Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A., Hepler,W.T.
            and Henderson,R.A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0173032-A 110 04-OCT-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..3410
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      667 a   1014 c    945 g    783 t      1 others
ORIGIN

  Query Match             100.0%;  Score 3409.6;  DB 6;  Length 3410;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy       2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy       2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy       2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy       2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700

Qy       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAA 3410
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAA 3410
```

RESULT 5
AX429961

```
LOCUS       AX429961                3410 bp    DNA     linear   PAT 21-JUN-2002
DEFINITION  Sequence 100 from Patent WO0198339.
ACCESSION   AX429961
VERSION     AX429961.1  GI:21541123
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Reed,S.G., Xu,J., Dillon,D.C., Retter,M.W. and Harlocker,S.L.
  TITLE     Compositions and methods for the therapy and diagnosis of breast
            cancer
  JOURNAL   Patent: WO 0198339-A 100 27-DEC-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..3410
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      667 a   1014 c    945 g    783 t      1 others
ORIGIN

  Query Match             100.0%;  Score 3409.6;  DB 6;  Length 3410;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
```

```
Qy       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy       2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy       2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy       2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy       2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCA 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCA 2700

Qy       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
```

```
RESULT 6
AY033593
LOCUS       AY033593                3410 bp    mRNA    linear   PRI 23-MAY-2001
DEFINITION  Homo sapiens prostein mRNA, complete cds.
ACCESSION   AY033593
VERSION     AY033593.1  GI:14192791
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3410)
  AUTHORS   Xu,J., Kalos,M., Stolk,J.A., Zasloff,E.J., Zhang,X., Houghton,R.L.,
            Filho,A.M., Nolasco,M., Badaro,R. and Reed,S.G.
  TITLE     Identification and characterization of prostein, a novel
            prostate-specific protein
  JOURNAL   Cancer Res. 61 (4), 1563-1568 (2001)
  MEDLINE   21139094
   PUBMED   11245466
REFERENCE   2  (bases 1 to 3410)
  AUTHORS   Xu,J., Kalos,M., Stolk,J.A., Zasloff,E.J., Zhang,X., Houghton,R.L.,
            Filho,A.M., Nolasco,M., Badaro,R. and Reed,S.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-APR-2001) Antigen Discovery, Corixa Corporation, 1124
            Columbia Street, Seattle, WA 98104, USA
FEATURES             Location/Qualifiers
     source          1..3410
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /tissue_type="prostate"
     CDS             284..1945
                     /codon_start=1
                     /product="prostein"
                     /protein_id="AAK54386.1"
                     /db_xref="GI:14192792"
                     /translation="MVQRLWVSRLLRHRKAQLLLVNLLTFGLEVCLAAGITYVPPLLL
```

```
                      EVGVEEKFMTMVLGIGPVLGLVCVPLLGSASDHWRGRYGRRRPFIWALSLGILLSLFL
                      IPRAGWLAGLLCPDPRPLELALLILGVGLLDFCGQVCFTPLEALLSDLFRDPDHCRQA
                      YSVYAFMISLGGCLGYLLPAIDWDTSALAPYLGTQEECLFGLLTLIFLTCVAATLLVA
                      EEAALGPTEPAEGLSAPSLSPHCCPCRARLAFRNLGALLPRLHQLCCRMPRTLRRLFV
                      AELCSWMALMTFTLFYTDFVGEGLYQGVPRAEPGTEARRHYDEGVRMGSLGLFLQCAI
                      SLVFSLVMDRLVQRFGTRAVYLASVAAFPVAAGATCLSHSVAVVTASAALTGFTFSAL
                      QILPYTLASLYHREKQVFLPKYRGDTGGASSEDSLMTSFLPGPKPGAPFPNGHVGAGG
                      SGLLPPPPALCGASACDVSVRVVVGEPTEARVVPGRGICLDLAILDSAFLLSQVAPSL
                      FMGSIVQLSQSVTAYMVSAAGLGLVAIYFATQVVFDKSDLAKYSA"
BASE COUNT     667 a   1014 c    945 g    783 t      1 others
ORIGIN

  Query Match            100.0%;  Score 3409.6;  DB 9;  Length 3410;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 3410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60

Qy         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAG 120

Qy        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTT 180

Qy        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTG 240

Qy        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTG 300

Qy        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT 360

Qy        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGT 420

Qy        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGT 480

Qy        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCG 540

Qy        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGC 600

Qy        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCAT 660

Qy        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCT 720

Qy        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTT 780

Qy        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAG 840

Qy        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900

Qy        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960

Qy        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020

Qy       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080

Qy       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140

Qy       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200

Qy       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260

Qy       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320

Qy       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380

Qy       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGG 1440

Qy       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGA 1500

Qy       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAG 1560

Qy       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGT 1620

Qy       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTG 1680

Qy       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCG 1740

Qy       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCC 1800

Qy       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTC 1860

Qy       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAG 1920

Qy       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1921 CGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCT 1980

Qy       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGT 2040

Qy       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTA 2100

Qy       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTG 2160

Qy       2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 ACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCC 2220

Qy       2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 ATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGC 2280

Qy       2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTG 2340

Qy       2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAG 2400

Qy       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGA 2460

Qy       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCA 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCA 2700

Qy       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATT 3240

Qy       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACA 3300

Qy       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3360

Qy       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 AAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAA 3410

RESULT 7
AX327336
LOCUS       AX327336                3320 bp    DNA     linear   PAT 07-JAN-2002
DEFINITION  Sequence 1 from Patent WO0181577.
ACCESSION   AX327336
VERSION     AX327336.1  GI:18097882
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lau,T., Lin,R.J., Parkes,D., Parry,G., Schneider,D.W.,
            Steinbrecher,R., van Heuit,P.T. and Wu,J.
  TITLE     Dna encoding the prost 03 polypeptide
  JOURNAL   Patent: WO 0181577-A 1 01-NOV-2001;
            SCHERING AKTIENGESELLSCHAFT (DE)
FEATURES             Location/Qualifiers
     source          1..3320
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             282..1943
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD20424.1"
                     /db_xref="GI:18097883"
                     /translation="MVQRLWVSRLLRHRKAQLLLVNLLTFGLEVCLAAGITYVPPLLL
                     EVGVEEKFMTMVLGIGPVLGLVCVPLLGSASDHWRGRYGRRRPFIWALSLGILLSLFL
                     IPRAGWLAGLLCPDPRPLELALLILGVGLLDFCGQVCFTPLEALLSDLFRDPDHCRQA
                     YSVYAFMISLGGCLGYLLPAIDWDTSALAPYLGTQEECLFGLLTLIFLTCVAATLLVA
                     EEAALGPTEPAEGLSAPSLSPHCCPCRARLAFRNLGALLPRLHQLCCRMPRTLRRLFV
                     AELCSWMALMTFTLFYTDFVGEGLYQGVPRAEPGTEARRHYDEGVRMGSLGLFLQCAI
                     SLVFSLVMDRLVQRFGTRAVYLASVAAFPVAAGATCLSHSVAVVTASAALTGFTFSAL
                     QILPYTLASLYHREKQVFLPKYRGDTGGASSEDSLMTSFLPGPKPGAPFPNGHVGAGG
                     SGLLPPPPALCGASACDVSVRVVVGEPTEARVVPGRGICLDLAILDSAFLLSQVAPSL
                     FMGSIVQLSQSVTAYMVSAAGLGLVAIYFATQVVFDKSDLAKYSA"
BASE COUNT      585 a   1013 c    944 g    778 t
ORIGIN

  Query Match             96.6%;  Score 3292.4;  DB 6;  Length 3320;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3316;  Conservative    0;  Mismatches    1;  Indels    3;  Gaps    2;

Qy          3 GAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGT 62
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGT 60
```

```
Qy         63 GATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAA 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAA 120

Qy        123 GCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCCTGGCTGATTCCTAGGCAGTTGG 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCCTGGCTGATTCCTAGGCAGTTGG 180

Qy        183 CGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGA 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGA 240

Qy        243 GTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGG 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGG 300

Qy        303 TGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTG 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTG 360

Qy        363 GCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGG 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGG 420

Qy        423 GGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCT 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCT 480

Qy        483 GTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGC 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGC 540

Qy        543 CCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCG 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCG 600

Qy        603 GCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCC 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCC 660

Qy        663 TGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGC 722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGC 720

Qy        723 TCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCA 782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCA 780

Qy        783 TGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTG 842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTG 840

Qy        843 CCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCT 902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCT 900

Qy        903 TCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCG 962
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCG 960

Qy        963 AGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCC 1022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCC 1020

Qy       1023 GCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCA 1082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCA 1080

Qy       1083 TGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGA 1142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGA 1140

Qy       1143 CCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAG 1202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAG 1200

Qy       1203 CTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGG 1262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGG 1260

Qy       1263 GGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGC 1322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGC 1320

Qy       1323 AGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCG 1382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCG 1380

Qy       1383 GTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGT 1442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGT 1440

Qy       1443 TCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGA 1502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGA 1500

Qy       1503 AGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCC 1562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCC 1560

Qy       1563 TGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGG 1622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGG 1620

Qy       1623 GTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTG 1682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTG 1680

Qy       1683 ATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGG 1742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 ATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGG 1740

Qy       1743 GCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCAT 1802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCAT 1800

Qy       1803 CCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTG 1862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTG 1860

Qy       1863 CCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCG 1922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCG 1920

Qy       1923 ACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCA 1982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 ACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCA 1980

Qy       1983 CTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTT 2042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTT 2040

Qy       2043 CTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGC 2102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 CTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGC 2100

Qy       2103 TGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGAC 2162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGAC 2160

Qy       2163 TGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCAT 2222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 TGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCAT 2220

Qy       2223 GCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCT 2282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      2221 GCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCT 2280

Qy      2283 CCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGT 2342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2281 CCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGT 2340

Qy      2343 TTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTT 2402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2341 TTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTT 2400

Qy      2403 TCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAGA 2460
             |||||||||||||||||||||||||||||||||||||||  | |||||||||||||||||
Db      2401 TCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGAAAGTTATTTGTAGGGGAAGA 2460

Qy      2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2461 GTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCT 2520

Qy      2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2521 GATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCA 2580

Qy      2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2581 CAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCAT 2640

Qy      2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2641 TGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCA 2700

Qy      2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2701 GGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGT 2760

Qy      2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2761 CTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAAT 2820

Qy      2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2821 TCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGT 2880

Qy      2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2881 CTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA 2940

Qy      2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2941 CTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACC 3000

Qy      3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3001 CCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACT 3060

Qy      3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3061 GCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGT 3120

Qy      3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3121 ATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGG 3180

Qy      3181 GAGGTCTTATCTCTCA-GGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTAT 3239
             |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db      3181 GAGGTCTTATCTCTCAGGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTAT 3240

Qy      3240 TTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGAC 3299
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3241 TTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGAC 3300

Qy      3300 AAAATTAAAGGCTTTCTTAT 3319
             ||||||||||||||||||||
Db      3301 AAAATTAAAGGCTTTCTTAT 3320
```

```
RESULT 8
AB060851
LOCUS       AB060851                3514 bp    mRNA    linear   PRI 13-JUN-2001
DEFINITION  Macaca fascicularis brain cDNA clone:QtrA-11310, full insert
            sequence.
ACCESSION   AB060851
VERSION     AB060851.1  GI:13874496
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Macaca fascicularis adult male temporal lobe right cDNA to mRNA,
            clone_lib:macaque brain cDNA library QtrA clone:QtrA-11310.
  ORGANISM  Macaca fascicularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (sites)
  AUTHORS   Osada,N., Hida,M., Kusuda,J., Tanuma,R., Iseki,K., Hirai,M.,
            Terao,K., Suzuki,Y., Sugano,S. and Hashimoto,K.
  TITLE     Isolation of full-length cDNA clones from macaque brain cDNA
            libraries
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3514)
  AUTHORS   Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-APR-2001) Katsuyuki Hashimoto, National Institute of
            Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
            1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
            (E-mail:khashi@nih.go.jp, URL:http://www.nih.go.jp/yoken/genebank/,
            Tel:81-3-5285-1111(ex.2120), Fax:81-3-5285-1181)
COMMENT     Lab host:    TOP10
            Vector:      pME18S-FL3 (Acc.No. AB009864)
            R.  Site1:   DraIII (CACTGTGTG)
            R.  Site2:   DraIII (CACCATGTG)
            Description: 1st strand cDNA was primed with an oligo(dT) primer
            [ATGTGGCCTTTTTTTTTTTTTTTTT]; double-stranded cDNA was synthesized
            using specific 5' and 3' primers and amplified by PCR.  The PCR
            product was digested with SfiI and size selection was performed to
            exclude fragments <1.5kb.The SfiI-digested PCR product was cloned
            into distinct DraIII sites of pME18S-FL3. XhoI sites just outside
            the DraIII sites can be used to isolate the cDNA insert.  Libraries
            were constructed by oligo-capping method
            (Sugano et al., , Institute of Medical Science, University of
            Tokyo).
            Custom primer used for sequencing
            ( 5' end primer [CTTCTGCTCTAAAAGCTGCG];
            3' end primer [CGACCTGCAGCTCGAGCACA] ).
FEATURES             Location/Qualifiers
     source          1..3514
                     /organism="Macaca fascicularis"
                     /db_xref="taxon:9541"
                     /clone="QtrA-11310"
                     /sex="male"
                     /tissue_type="temporal lobe right"
                     /clone_lib="macaque brain cDNA library QtrA"
                     /dev_stage="adult"
     CDS             454..2115
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAB46871.1"
                     /db_xref="GI:13874497"
                     /translation="MVQRLWVSRLLRHRKAQLLLINLLTFGLEVCLAAGITYVPPLLL
                     EVGVEEKFMTMVLGIGPVLGLVSVPLLGSASDHWRGRYGRRRPFIWALSLGILLSLFL
                     IPRAGWLAGLLCPDPRPLELALLILGVGLLDFCGQVCFTPLEALLSDLFRDPDHCRQA
                     YSVYAFMISLGGCLGYLLPAIDWDTSALAPYLGTQEECLFGLLTLIFLTCVAATLLVA
                     EEAALGPAEPAEGLSAPSLPSHCCPCWARLAFRNLGALLPRLHQLCCRMPRTLRRLFV
                     AELCSWMALMTFTLFYTDFVGEGLYQGVPRAELGTEARRHYDEGVRMGSLGLFLQCAI
                     SLVFSLVMDRLVQRFGTRAVYLASVAAFPVAAGATCLSHSVAVVTASAALTGFTFSAL
                     QILPYTLASLYHRERQVFLPKYRGDAGGTSSEDSLMTSFLPGPKPGAPFPNGHVGAGG
                     SGLLPPPPALCGASACDVSVRVVVGEPTEARVVPGRGICLDLAILDSAFLLSQVAPSL
                     FMGSIVQLSQSVTAYMVSAAGLGLVAIYFATQVVFDKSDLAKYSV"
BASE COUNT      647 a   1039 c   1003 g    825 t
ORIGIN
```

```
Query Match             85.0%;  Score 2900.2;  DB 9;  Length 3514;
Best Local Similarity   94.8%;  Pred. No. 0;
Matches 3155;  Conservative    0;  Mismatches  118;  Indels   54;  Gaps   13;

Qy         51 CAGGATCTGAGTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGC 110
Db        221 CAGGATCTGAGTGATGAGATGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGC 280

Qy        111 ATGGGCTGAGAAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTC 170
Db        281 ATGGGCTGAGAAGCTGGACCGGCACCAAAGGGCTGGCGGAAATGGGCGCCTGGCTGATTC 340

Qy        171 CTAGGCAGTTGGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGA 230
Db        341 CTAGGCAGTTGGCGGCAGCAAGGAGGAGAGGCCGTGGCTTCCGGAGCAGAGCGGAGACGA 400

Qy        231 AGCAGTTCTGGAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCC 290
Db        401 AGCAGTTCTGGAGTGCTTAAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCC 460

Qy        291 AGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACC 350
Db        461 AGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGTTGATCAACC 520

Qy        351 TGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGC 410
Db        521 TGCTAACCTTTGGCCTGGAAGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGC 580

Qy        411 TGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGC 470
Db        581 TGCTGGAAGTGGGGGTAGAAGAGAAGTTCATGACCATGGTGCTGGGCATCGGTCCAGTGC 640

Qy        471 TGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATG 530
Db        641 TGGGCCTGGTCTCTGTCCCACTCCTAGGCTCAGCCAGTGACCACTGGCGCGGACGCTATG 700

Qy        531 GCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCA 590
Db        701 GCCGCCGGCGGCCCTTCATCTGGGCGCTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCA 760

Qy        591 TCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGG 650
Db        761 TCCCAAGGGCTGGCTGGCTGGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGG 820

Qy        651 CACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCAC 710
Db        821 CACTGCTCATCCTGGGCGTAGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCAC 880

Qy        711 TGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTG 770
Db        881 TGGAGGCCCTGCTCTCTGACCTGTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCCG 940

Qy        771 TCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACT 830
Db        941 TCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACT 1000

Qy        831 GGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGC 890
Db       1001 GGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGC 1060

Qy        891 TCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGC 950
Db       1061 TCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCCGAGGAGGCAGCAC 1120

Qy        951 TGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTC 1010
Db       1121 TGGGCCCCGCCGAGCCAGCGGAAGGGCTGTCTGCCCCCTCCCTGCCGTCCCACTGCTGTC 1180

Qy       1011 CATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGC 1070
Db       1181 CGTGCTGGGCCCGCCTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGC 1240

Qy       1071 TGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGA 1130
Db       1241 TGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGA 1300

Qy       1131 TGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGG 1190
Db       1301 TGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTATACCAGG 1360

Qy       1191 GCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGA 1250
Db       1361 GCGTGCCCAGAGCTGAGCTGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGA 1420

Qy       1251 TGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGG 1310
Db       1421 TGGGCAGTCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGG 1480

Qy       1311 ACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCC 1370
Db       1481 ACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATCTGGCCAGTGTGGCAGCTTTCC 1540

Qy       1371 CTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCG 1430
Db       1541 CTGTGGCTGCCGGTGCCACGTGCCTGTCCCACAGTGTGGCTGTGGTGACGGCTTCAGCCG 1600

Qy       1431 CCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCT 1490
Db       1601 CCCTCACTGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCT 1660

Qy       1491 ACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCA 1550
Db       1661 ACCACCGGGAGAGGCAGGTGTTCCTGCCCAAATACCGAGGGGACGCTGGAGGCACTAGCA 1720

Qy       1551 GTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTA 1610
Db       1721 GTGAGGACAGCCTGATGACTAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTA 1780

Qy       1611 ATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGG 1670
Db       1781 ATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTTCCACCTCCACCCGCGCTCTGCGGGG 1840

Qy       1671 CCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGG 1730
Db       1841 CCTCTGCCTGCGATGTCTCTGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGG 1900

Qy       1731 TTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCC 1790
Db       1901 TTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCC 1960

Qy       1791 AGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCT 1850
Db       1961 AGGTGGCCCCGTCCCTGTTCATGGGCTCCATCGTCCAGCTCAGCCAGTCTGTCACTGCCT 2020

Qy       1851 ATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTAT 1910
Db       2021 ATATGGTGTCTGCTGCAGGCCTGGGTCTGGTTGCCATTTACTTTGCTACACAGGTAGTAT 2080

Qy       1911 TTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGA 1970
Db       2081 TTGACAAGAGCGACTTGGCCAAATACTCGGTGTAGAAAACTTCCAGCACATTGGGGTGGA 2140

Qy       1971 GGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTC-CTGTTAGCCCCATGGGGCTGCCGGGC 2029
Db       2141 GGGCCTGCCTCACTGGGTCCCAGCTCCCCACTCTTTGTTAGCCCCATGGGGCTGCTGGGC 2200

Qy       2030 TGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGC 2089
Db       2201 TGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGC 2260

Qy       2090 TGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCT 2149
Db       2261 TGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCATCCCTCTCCCTCCTCCCCAGTCTCT 2320

Qy       2150 AGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCC 2209
```

```
              ||||||||||||||||| |||||||||||||||||||||||||||| |||||||||||
Db       2321 AGGGCTGCCTGACTGGAAGCCTTCCAAGGGGGTTTCAGTCTGGACTTCTACAGGGAGGCT 2380

Qy       2210 AGAAGG----------------GCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGAT 2253
              ||||||                |||||||||||||||||| ||||||||||||||||||
Db       2381 AGAAGGGCAGGGCATTTGATTCGCTCCATGCACTGGAATGTGGGGACTCTGCAGGTGGAT 2440

Qy       2254 TACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTT 2313
              |||||||||||||||||||||||||||||||| | |||||| ||||||||||||| |||
Db       2441 TACCCAGGCTCAGGGTTAACAGCTAGCCTCCTGGCTGAGACATACCTAGAGAAGGGGTTT 2500

Qy       2314 TGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAG-CCCCTTAACCTGCAGC 2372
              ||||||||| ||||||||||||||||||||||| ||||||| ||||||||||||||||
Db       2501 TGGGAGCTGAGTAAACTCAGTCACCTGGTTTCCCACCTCTAAGCCCCCTTAACCTGCAGC 2560

Qy       2373 TTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTT 2432
              ||| ||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db       2561 TTCATTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCGTGGGATTT 2620

Qy       2433 GAACATATG--ACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGT 2490
              |||| ||||  | |||||||||||||||||||||||||||||||||||      |||||||
Db       2621 GAACGTATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACAC------ACCAGGT 2674

Qy       2491 CCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGA 2550
              |||||||||||||||||||| |||||||||||||| ||  ||||||||||||||||||||
Db       2675 CCCCTCAGCCCACAGCACTGCCTTTTTGCTGATCC-CCTGACTCTTACCTTTTATCAGGA 2733

Qy       2551 TGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTT 2610
               ||||||| ||||||| | ||||||||||| || |||||||||||||||||||||||||
Db       2734 CGTGGCCTATTGGTCCCTTTGTTGCCATCATAGGGACACAGGCATTTAAATATTTAACTT 2793

Qy       2611 ATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATAT 2670
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||| ||
Db       2794 ATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTGTGTGTTGGTGTCTAAGAT 2853

Qy       2671 TTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTG 2730
              |||||||||||  |||||||||||||||||||||| ||||||| ||||||||||||||||
Db       2854 TTGGGTAGGGT--GGGATCCCCAACAATCAGGTCCACTGAGATCACTGGTCATTGGGCTG 2911

Qy       2731 ATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACC 2790
              |||||||||||||||||||||||||||||||||||| || ||||||||||||||||||||
Db       2912 ATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCTCCTCAAAATGCCTAACCCAGGACC 2971

Qy       2791 TTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGT 2850
              |||||||| |||||||||| | ||||||||||||||||||||||||||||||||||| ||
Db       2972 TTGGAAATTTACTCATCCCGACTGATAATTCCAAATGCTGTTACCCAAGGTTAGGGGGT 3031

Qy       2851 TGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCT 2910
              |||||||||| |||||||||||||||||||||||||| ||||||||||||||||| ||||
Db       3032 TGAAGGAAGGTGGAGGGTGGGGCTTCAGGTCTCAACAGCTTCCCTAACCACCCCTTTTCT 3091

Qy       2911 CTTGGCCCAGCCTGGTTCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTG 2970
              |||||||||||||||||||||||||||| ||||||||||||| |||||||||||||||||
Db       3092 CTTGGCCCAGCCTGGTTCCCCCACTTCTACTCCCCTCTACTGTCTCTAGGACTGGGCTG 3151

Qy       2971 ATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCC 3030
              ||||||||||||| ||||||                 ||| ||||||||||||||||
Db       3152 ATGAAGGCACTGCCTGAAATTT-----------------CCCTCACCCCCAACTTTCCCC 3194

Qy       3031 ACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCC 3090
              ||  |||||||||||||||||||||||| ||||||||||||||||||||||| |||||||
Db       3195 ACTGGCTCCACAACCCTGTTTGGAGCTGTTGCAGGACCAGAAGCACAAAGTGTGGTTTCC 3254

Qy       3091 CAAGCCTTTGTCCATCTCAG-CCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGA 3149
              || |||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db       3255 CAGGCCTTTGTCCATCTCAGCCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGA 3314

Qy       3150 AACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAG 3209
              |||||||||||||||||||||||||||| |||||||||||||||||| ||||||||||||
Db       3315 AACTCAGGAGCACCCCCTGCCTGAGCTAA-GGAGGTCTTATCTCTCA-GGGGGGTTTAAG 3372

Qy       3210 TGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGT 3269
              ||||||||||||||||   |||||||||||||||||||||  ||||||||||||||||||
Db       3373 TGCCGTTTGCAATAAT---GTCTTATTTATTTAGCGGGGCAAATATTTTATACTGTAAGT 3429

Qy       3270 GAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAA 3329
              ||||||||  |||||||||||||||||||| ||||||||||||||||||||||||||||
Db       3430 GAGCAATC--AGTATAATGTTTATGGTGATGAAATTAAAGGCTTTCTTATATGTTTAAAA 3487

Qy       3330 AAAAAAAAAAAAAAAAAAAAAAAAAAA 3356
              |||||||||||||||||||||||||||
Db       3488 AAAAAAAAAAAAAAAAAAAAAAAAAAA 3514
```

RESULT 9
AX200995
```
LOCUS       AX200995                4034 bp    DNA     linear   PAT 29-AUG-2001
DEFINITION  Sequence 625 from Patent WO0151633.
ACCESSION   AX200995
VERSION     AX200995.1  GI:15390822
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4034)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Fanger,G.R., Day,C.H., Retter,M.W.,
            Stolk,J.A., Skeiky,Y.A., Wang,A. and Meagher,M.J.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0151633-A 625 19-JUL-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..4034
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      721 a   1226 c   1141 g    946 t
ORIGIN

  Query Match             75.8%;  Score 2585.4;  DB 6;  Length 4034;
  Best Local Similarity   82.4%;  Pred. No. 0;
  Matches 3325;  Conservative    0;  Mismatches    1;  Indels  709;  Gaps    3;

Qy          4 AACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTG 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTG 60

Qy         64 ATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAG 120

Qy        124 CTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGC 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGC 180

Qy        184 GGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAG 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAG 240

Qy        244 TGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGT 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGT 300

Qy        304 GAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGG 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGG 360

Qy        364 CCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGG 420

Qy        424 GGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTG 480
```

```
Qy    484 TGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC 543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC 540

Qy    544 CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGG 603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGG 600

Qy    604 CTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCT 663
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCT 660

Qy    664 GGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT 723
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT 720

Qy    724 CTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCAT 783
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCAT 780

Qy    784 GATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGC 843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGC 840

Qy    844 CCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTT 903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTT 900

Qy    904 CCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGA 963
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGA 960

Qy    964 GCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCG 1023
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCG 1020

Qy   1024 CTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCAT 1083
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCAT 1080

Qy   1084 GCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGAC 1143
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGAC 1140

Qy   1144 CTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGC 1203
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGC 1200

Qy   1204 TGAGCCGGGCACCGAGGCCCGGAGACACTATGATGA------------------------ 1239
          ||||||||||||||||||||||||||||||||||||
Db   1201 TGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGTAAGGCCTTGGCAGCCAGCAG 1260

Qy   1240 ------------------------------------------------------------ 1239

Db   1261 AGGCTGGTGTGGGAGCCGCCCACCAGAGACGACACTCGGGGCTGTGTCTGGGCTGGTGCC 1320

Qy   1240 ------------------------------------------------------------ 1239

Db   1321 TCTCCATCCTGGCCCCGACTTCTCTGTCAGGAAAGTGGGGATGGACCCCATCTGCATACA 1380

Qy   1240 ------------------------------------------------------------ 1239

Db   1381 CGGCTTCTCATGGGTGTGGAACATCTCTGCTTGCGGTTTCAGGAAGGCCTCTGGCTGCTC 1440

Qy   1240 ------------------------------------------------------------ 1239

Db   1441 TAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGACAGAAGGAAAGGCGGAGCTTATTCAA 1500

Qy   1240 ------------------------------------------------------------ 1239

Db   1501 AGTCTAGAGGGAGTGGAGGAGTTAAGGCTGGATTTCAGATCTGCCTGGTTCCAGCCGCAG 1560

Qy   1240 ------------------------------------------------------------ 1239

Db   1561 TGTGCCCTCTGCTCCCCCAACGACTTTCCAAATAATCTCACCAGCGCCTTCCAGCTCAGG 1620

Qy   1240 ------------------------------------------------------------ 1239

Db   1621 CGTCCTAGAAGCGTCTTGAAGCCTATGGCCAGCTGTCTTTGTGTTCCCTCTCACCCGCCT 1680

Qy   1240 ------------------------------------------------------------ 1239

Db   1681 GTCCTCACAGCTGAGACTCCCAGGAAACCTTCAGACTACCTTCCTCTGCCTTCAGCAAGG 1740

Qy   1240 ------------------------------------------------------------ 1239

Db   1741 GGCGTTGCCCACATTCTCTGAGGGTCAGTGGAAGAACCTAGACTCCCATTGCTAGAGGTA 1800

Qy   1240 ------------------------------------------------------------ 1239

Db   1801 GAAAGGGGAAGGGTGCTGGGGAGCAGGGCTGGTCCACAGCAGGTCTCGTGCAGCAGGTAC 1860

Qy   1240 ------------------------------------------------------------ 1239

Db   1861 CTGTGGTTCCGCCTTCTCATCTCCCTGAGACTGCTCCGACCCTTCCCTCCCAGGCTCTGT 1920

Qy   1240 ----------------------AGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGT 1277
                                ||||||||||||||||||||||||||||||||||||||
Db   1921 CTGATGGCCCCTCTCCCTCTGCAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGT 1980

Qy   1278 GCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTC 1337
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 GCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTC 2040

Qy   1338 GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT 1397
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT 2100

Qy   1398 CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCC 1457
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCC 2160

Qy   1458 TGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGC 1517
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 TGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGC 2220

Qy   1518 CCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCC 1577
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 CCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCC 2280

Qy   1578 TGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTG 1637
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 TGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTG 2340

Qy   1638 GCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTG 1697
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTG 2400

Qy   1698 TGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACC 1757
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACC 2460

Qy   1758 TCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCT 1817
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 TCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCT 2520

Qy   1818 CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTC 1877
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTC 2580

Qy   1878 TGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT 1937
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 TGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT 2640

Qy   1938 CAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCC 1997
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      2641 CAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCC 2700

Qy      1998 CCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAG 2057
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2701 CCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAG 2760

Qy      2058 TAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2117
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2761 TAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2820

Qy      2118 TGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAG 2177
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2821 TGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAG 2880

Qy      2178 GGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGG 2237
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2881 GGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGG 2940

Qy      2238 GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACA 2297
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2941 GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACA 3000

Qy      2298 CCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGC 2357
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3001 CCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGC 3060

Qy      2358 CCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACAC 2417
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3061 CCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACAC 3120

Qy      2418 TCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAAC 2475
             ||||||||||||||||||||||||   | |||||||||||||||||||||||||||||||
Db      3121 TCCTCCATGGGATTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAAC 3180

Qy      2476 ACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCT 2535
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3181 ACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCT 3240

Qy      2536 TACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCAT 2595
             |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db      3241 TACCTTTTATCAGGATGT-GCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCAT 3299

Qy      2596 TTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGT 2655
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3300 TTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGT 3359

Qy      2656 GTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAG 2715
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3360 GTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAG 3419

Qy      2716 CTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAAT 2775
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3420 CTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAAT 3479

Qy      2776 GCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTAC 2835
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3480 GCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTAC 3539

Qy      2836 CCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT 2895
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3540 CCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT 3599

Qy      2896 AACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTC 2955
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3600 AACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTC 3659

Qy      2956 TCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTA 3015
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3660 TCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTA 3719

Qy      3016 CCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCA 3075
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3720 CCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCA 3779

Qy      3076 CAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGG 3135
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3780 CAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGG 3839

Qy      3136 GAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTC 3195
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3840 GAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTC 3899

Qy      3196 AGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATAT 3255
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3900 AGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATAT 3959

Qy      3256 TTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTC 3315
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3960 TTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTC 4019

Qy      3316 TTATATGTTTAAAAA 3330
             |||||||||||||||
Db      4020 TTATATGTTTAAAAA 4034
```

```
RESULT 10
AX267730
LOCUS       AX267730                4034 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 704 from Patent WO0173032.
ACCESSION   AX267730
VERSION     AX267730.1  GI:16516402
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
            Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A., Hepler,W.T.
            and Henderson,R.A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0173032-A 704 04-OCT-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..4034
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      721 a   1226 c   1141 g    946 t
ORIGIN

  Query Match             75.8%;  Score 2585.4;  DB 6;  Length 4034;
  Best Local Similarity   82.4%;  Pred. No. 0;
  Matches 3325;  Conservative    0;  Mismatches    1;  Indels  709;  Gaps    3;

Qy          4 AACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTG 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTG 60

Qy         64 ATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAG 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAG 120

Qy        124 CTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGC 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGC 180

Qy        184 GGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAG 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAG 240

Qy        244 TGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGT 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGT 300
```

```
Qy        304 GAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGG 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGG 360

Qy        364 CCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGG 420

Qy        424 GGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTG 480

Qy        484 TGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC 540

Qy        544 CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGG 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGG 600

Qy        604 CTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCT 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCT 660

Qy        664 GGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT 720

Qy        724 CTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCAT 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCAT 780

Qy        784 GATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGC 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGC 840

Qy        844 CCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTT 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTT 900

Qy        904 CCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGA 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGA 960

Qy        964 GCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCG 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCG 1020

Qy       1024 CTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCAT 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCAT 1080

Qy       1084 GCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGAC 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGAC 1140

Qy       1144 CTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGC 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGC 1200

Qy       1204 TGAGCCGGGCACCGAGGCCCGGAGACACTATGATGA------------------------ 1239
              ||||||||||||||||||||||||||||||||||||
Db       1201 TGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGTAAGGCCTTGGCAGCCAGCAG 1260

Qy       1240 ------------------------------------------------------------ 1239

Db       1261 AGGCTGGTGTGGGAGCCGCCCACCAGAGACGACACTCGGGGCTGTGTCTGGGCTGGTGCC 1320

Qy       1240 ------------------------------------------------------------ 1239

Db       1321 TCTCCATCCTGGCCCCGACTTCTCTGTCAGGAAAGTGGGGATGGACCCCATCTGCATACA 1380

Qy       1240 ------------------------------------------------------------ 1239

Db       1381 CGGCTTCTCATGGGTGTGGAACATCTCTGCTTGCGGTTTCAGGAAGGCCTCTGGCTGCTC 1440

Qy       1240 ------------------------------------------------------------ 1239

Db       1441 TAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGACAGAAGGAAAGGCGGAGCTTATTCAA 1500

Qy       1240 ------------------------------------------------------------ 1239

Db       1501 AGTCTAGAGGGAGTGGAGGAGTTAAGGCTGGATTTCAGATCTGCCTGGTTCCAGCCGCAG 1560

Qy       1240 ------------------------------------------------------------ 1239

Db       1561 TGTGCCCTCTGCTCCCCCAACGACTTTCCAAATAATCTCACCAGCGCCTTCCAGCTCAGG 1620

Qy       1240 ------------------------------------------------------------ 1239

Db       1621 CGTCCTAGAAGCGTCTTGAAGCCTATGGCCAGCTGTCTTTGTGTTCCCTCTCACCCGCCT 1680

Qy       1240 ------------------------------------------------------------ 1239

Db       1681 GTCCTCACAGCTGAGACTCCCAGGAAACCTTCAGACTACCTTCCTCTGCCTTCAGCAAGG 1740

Qy       1240 ------------------------------------------------------------ 1239

Db       1741 GGCGTTGCCCACATTCTCTGAGGGTCAGTGGAAGAACCTAGACTCCCATTGCTAGAGGTA 1800

Qy       1240 ------------------------------------------------------------ 1239

Db       1801 GAAAGGGGAAGGGTGCTGGGGAGCAGGGCTGGTCCACAGCAGGTCTCGTGCAGCAGGTAC 1860

Qy       1240 ------------------------------------------------------------ 1239

Db       1861 CTGTGGTTCCGCCTTCTCATCTCCCTGAGACTGCTCCGACCCTTCCCTCCCAGGCTCTGT 1920

Qy       1240 ----------------------AGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGT 1277
                                    ||||||||||||||||||||||||||||||||||||||
Db       1921 CTGATGGCCCCTCTCCCTCTGCAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGT 1980

Qy       1278 GCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTC 1337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTC 2040

Qy       1338 GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT 1397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT 2100

Qy       1398 CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCC 1457
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCC 2160

Qy       1458 TGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGC 1517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 TGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGC 2220

Qy       1518 CCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCC 1577
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCC 2280

Qy       1578 TGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTG 1637
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 TGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTG 2340

Qy       1638 GCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTG 1697
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTG 2400

Qy       1698 TGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACC 1757
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACC 2460

Qy       1758 TCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCATCCCTGTTTATGGGCT 1817
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     2461 TCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCT 2520

Qy     1818 CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTC 1877
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2521 CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTC 2580

Qy     1878 TGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT 1937
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2581 TGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT 2640

Qy     1938 CAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCC 1997
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2641 CAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCC 2700

Qy     1998 CCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAG 2057
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2701 CCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAG 2760

Qy     2058 TAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2117
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2761 TAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2820

Qy     2118 TGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAG 2177
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2821 TGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAG 2880

Qy     2178 GGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGG 2237
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2881 GGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGG 2940

Qy     2238 GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACA 2297
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2941 GACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACA 3000

Qy     2298 CCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGC 2357
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3001 CCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGC 3060

Qy     2358 CCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACAC 2417
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3061 CCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACAC 3120

Qy     2418 TCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAAC 2475
            ||||||||||||||||||||||||    |||||||||||||||||||||||||||||||
Db     3121 TCCTCCATGGGATTTGAACATATGAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAAC 3180

Qy     2476 ACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCT 2535
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3181 ACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCT 3240

Qy     2536 TACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCAT 2595
            |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db     3241 TACCTTTTATCAGGATGT-GCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCAT 3299

Qy     2596 TTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGT 2655
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3300 TTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGT 3359

Qy     2656 GTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAG 2715
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3360 GTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAG 3419

Qy     2716 CTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAAT 2775
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3420 CTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAAT 3479

Qy     2776 GCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTAC 2835
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3480 GCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTAC 3539

Qy     2836 CCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT 2895
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3540 CCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT 3599

Qy     2896 AACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTC 2955
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3600 AACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTC 3659

Qy     2956 TCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTA 3015
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3660 TCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTA 3719

Qy     3016 CCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCA 3075
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3720 CCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCA 3779

Qy     3076 CAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGG 3135
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3780 CAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGG 3839

Qy     3136 GAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTC 3195
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3840 GAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTC 3899

Qy     3196 AGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATAT 3255
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3900 AGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATAT 3959

Qy     3256 TTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTC 3315
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3960 TTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTC 4019

Qy     3316 TTATATGTTTAAAAA 3330
            |||||||||||||||
Db     4020 TTATATGTTTAAAAA 4034
```

```
RESULT 11
AB062977
LOCUS       AB062977                2917 bp    mRNA    linear   PRI 13-JUN-2001
DEFINITION  Macaca fascicularis brain cDNA clone:QmoA-10594, full insert
            sequence.
ACCESSION   AB062977
VERSION     AB062977.1  GI:14388390
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Macaca fascicularis adult male medulla oblogata cDNA to mRNA,
            clone_lib:macaque brain cDNA library QmoA clone:QmoA-10594.
  ORGANISM  Macaca fascicularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (sites)
  AUTHORS   Osada,N., Hida,M., Kusuda,J., Tanuma,R., Iseki,K., Hirai,M.,
            Terao,K., Suzuki,Y., Sugano,S. and Hashimoto,K.
  TITLE     Isolation of full-length cDNA clones from macaque brain cDNA
            libraries
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2917)
  AUTHORS   Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JUN-2001) Katsuyuki Hashimoto, National Institute of
            Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
            1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
            (E-mail:khashi@nih.go.jp, URL:http://www.nih.go.jp/yoken/genebank/,
            Tel:81-3-5285-1111(ex.2120), Fax:81-3-5285-1181)
COMMENT     Lab.host:     TOP10
            Vector:       pME18S-FL3 (Acc.No. AB009864)
            R.  Site1:    DraIII (CACTGTGTG)
            R.  Site2:    DraIII (CACCATGTG)
            Description: 1st strand cDNA was primed with an oligo(dT) primer
            [ATGTGGCCTTTTTTTTTTTTTTTTT]; double-stranded cDNA was synthesized
            using specific 5' and 3' primers and amplified by PCR.  The PCR
            product was digested with SfiI and size selection was performed to
            exclude fragments <1.5kb.The SfiI-digested PCR product was cloned
            into distinct DraIII sites of pME18S-FL3. XhoI sites just outside
            the DraIII sites can be used to isolate the cDNA insert.  Libraries
```

```
           were constructed by oligo-capping method
           (Sugano et al., , Institute of Medical Science, University of
           Tokyo).
           Custom primer used for sequencing
           ( 5' end primer [CTTCTGCTCTAAAAGCTGCG];
           3' end primer [CGACCTGCAGCTCGAGCACA] ).
FEATURES             Location/Qualifiers
     source          1. .2917
                     /organism="Macaca fascicularis"
                     /db_xref="taxon:9541"
                     /clone="QmoA-10594"
                     /sex="male"
                     /tissue_type="medulla oblogata"
                     /clone_lib="macaque brain cDNA library QmoA"
                     /dev_stage="adult"
     CDS             24. .1529
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAB60745.1"
                     /db_xref="GI:14388391"
                     /translation="MTMVLGIGPVLGLVSVPLLGSASDHWRGRYGRRRPFIWALSLGI
                     LLSLFLIPRAGWLAGLLCPDPRPLELALLILGVGLLDFCGQVCFTPLEALLSDLFRDP
                     DHCRQAYSVYTFMISLGGCLGYLLPAIDWDTSALAPYLGTQEECLFGLLTLIFLTCVA
                     ATLLVAEEAALGPAEPAEGLSAPSLPSHCCPCWARLAFRNLGALLPRLHQLCCRMPRT
                     LRRLFVAELCSWMALMTFTLFYTDFVGEGLYQGVPRAELGTEARRHYDEGVRMGSLGL
                     FLQCAISLVFSLVVDRLVQRFGTRAVYLASVAAFPVAAGATCLSHSVAVVTASAALTG
                     FTFSALQILPYTLASLYHRERQVFLPKYRGDAGGTSSEDSLMTSFLPGPKPGAPFPNG
                     HVGAGGSGLLPPPPALCGASACDVSRVVVGEPTEARVVPGRGICLDLAILDSAFLLS
                     QVAPSLFMGSIVQLSQSVTAYMVSAAGLGLVAIYFATQVVFDKSDLAKYSV"
BASE COUNT       514 a    868 c    815 g    720 t
ORIGIN

  Query Match             74.7%;  Score 2545.6;  DB 9;  Length 2917;
  Best Local Similarity   94.6%;  Pred. No. 0;
  Matches 2792;  Conservative    0;  Mismatches  104;  Indels   54;  Gaps   13;

Qy        417 AAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCC 476
Db          1 AAGTGGGGGTAGAAGAGAAGTTCATGACCATGGTGCTGGGCATCGGTCCAGTGCTGGGCC 60

Qy        477 TGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCC 536
Db         61 TGGTCTCTGTCCCACTCCTAGGCTCAGCCAGTGACCACTGGCGCGGACGCTATGGCCGCC 120

Qy        537 GCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAA 596
Db        121 GGCGGCCCTTCATCTGGGCGCTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAA 180

Qy        597 GGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGC 656
Db        181 GGGCTGGCTGGCTGGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGC 240

Qy        657 TCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGG 716
Db        241 TCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGG 300

Qy        717 CCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATG 776
Db        301 CCCTGCTCTCTGACCTGTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCCGTCTATA 360

Qy        777 CCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACA 836
Db        361 CCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACA 420

Qy        837 CCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCC 896
Db        421 CCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCC 480

Qy        897 TCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCC 956
Db        481 TCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCCGAGGAGGCAGCACTGGGCC 540

Qy        957 CCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCC 1016
Db        541 CCGCCGAGCCAGCGGAAGGGCTGTCTGCCCCCTCCCTGCCGTCCCACTGCTGTCCGTGCT 600

Qy       1017 GGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCT 1076
Db        601 GGGCCCGCCTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCT 660

Qy       1077 GCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCAC 1136
Db        661 GCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCAC 720

Qy       1137 TCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGC 1196
Db        721 TCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTATACCAGGGCGTGC 780

Qy       1197 CCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCA 1256
Db        781 CCAGAGCTGAGCTGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCA 840

Qy       1257 GCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGC 1316
Db        841 GTCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCGTGGACCGGC 900

Qy       1317 TGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGG 1376
Db        901 TGGTGCAGCGATTCGGCACTCGAGCAGTCTATCTGGCCAGTGTGGCAGCTTTCCCTGTGG 960

Qy       1377 CTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCA 1436
Db        961 CTGCCGGTGCCACGTGCCTGTCCCACAGTGTGGCTGTGGTGACGGCTTCAGCCGCCCTCA 1020

Qy       1437 CCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACC 1496
Db       1021 CTGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACC 1080

Qy       1497 GGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGG 1556
Db       1081 GGGAGAGGCAGGTGTTCCTGCCCAAATACCGAGGGGACGCTGGAGGCACTAGCAGTGAGG 1140

Qy       1557 ACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGAC 1616
Db       1141 ACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGAC 1200

Qy       1617 ACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTG 1676
Db       1201 ACGTGGGTGCTGGAGGCAGTGGCCTGCTTCCGCCTCCACCCGCGCTCTGCGGGGCCTCTG 1260

Qy       1677 CCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGG 1736
Db       1261 CCTGCGATGTCTCTGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGG 1320

Qy       1737 GCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGG 1796
Db       1321 GCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGG 1380

Qy       1797 CCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGG 1856
Db       1381 CCCCGTCCCTGTTCATGGGCTCCATCGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGG 1440

Qy       1857 TGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACA 1916
Db       1441 TGTCTGCTGCAGGCCTGGGTCTGGTTGCCATTTACTTTGCTACACAGGTAGTATTTGACA 1500

Qy       1917 AGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCT 1976
Db       1501 AGAGCGACTTGGCCAAATACTCGGTGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCT 1560

Qy       1977 GCCTCACTGGGTCCCAGCTCCCCGCTC-CTGTTAGCCCCATGGGGCTGCCGGGCTGGCCG 2035
Db       1561 GCCTCACTGGGTCCCAGCTCCCCACTCTTTGTTAGCCCCATGGGGCTGCTGGGCTGGCCG 1620

Qy       2036 CCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGT 2095
```

```
Db   1621 CCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGT 1680

Qy   2096 GCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCT 2155
Db   1681 GCGTAGCTGCACAGCTGGGGGCTGGGGCATCCCTCTCCCTTCTCCCCAGTCTCTAGGGCT 1740

Qy   2156 GCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGG 2215
Db   1741 GCCTGACTGGAAGCCTTCCAAGGGAGTTTCAGTCTGGACTTCTACAGGGAGGCTAGAAGG 1800

Qy   2216 ----------------GCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCA 2259
Db   1801 GCAGGGCATTTGATTCGCTCCATGCACTGGAATGTGGGGACTCTGCAGGTGGATTACCCA 1860

Qy   2260 GGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAG 2319
Db   1861 GGCTCAGGGTTAACAGCTAGCCTCCTGGCTGAGACATACCTAGAGAAGGGGTTTTGGGAG 1920

Qy   2320 CTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAG-CCCCTTAACCTGCAGCTTCGTT 2378
Db   1921 CTGAGTAAACTCAGTCACCTGGTTTCCCACCTCTAAGCCCCCTTAACCTGCAGCTTCGTT 1980

Qy   2379 TAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACAT 2438
Db   1981 TAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCGTGGGATTTGAACGT 2040

Qy   2439 ATG--ACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTC 2496
Db   2041 ATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACAC------ACCAGGTCCCCTC 2094

Qy   2497 AGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGC 2556
Db   2095 AGCCCACAGCACTGCCTTTTTGCTGATCC-CCCGACTCTTACCTTTTATCAGGACGTGGC 2153

Qy   2557 CTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTAT 2616
Db   2154 CTATTGGTCCCTTTGTTGCCATCACAGGGACACAGGCATTTAAATATTTAACTTATTTAT 2213

Qy   2617 TTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGT 2676
Db   2214 TTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAAGATTTGGGT 2273

Qy   2677 AGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATT 2736
Db   2274 AGGGT--GGGATCCCCAACAATCAGGTCCACTGAGATCACTGGTCATTGGGCTGATCATT 2331

Qy   2737 GCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAA 2796
Db   2332 GCCAGAATCTTCTTCTCCTGGGGTCTGGCTCCTCAAAATGCCTAACCCAGGACCTTGGAA 2391

Qy   2797 ATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGG 2856
Db   2392 ATTTTACTCATCCCGACTGATAATTCCAAATGCTGTTACCCAAGGTTAGGGGGTTGAAGG 2451

Qy   2857 AAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGC 2916
Db   2452 AAGGTGGAGGGTGGGGCTTCAGGTCTCAACAGCTTCCCTAACCACCCCTTTTCTCTTGGC 2511

Qy   2917 CCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTAGGACTGGGCTGATGAAG 2976
Db   2512 CCAGCCTGGTTCCCCCCACTTCTACTCCCCTCTACTGTCTCTAGGACTGGGCTGATGAAG 2571

Qy   2977 GCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGC 3036
Db   2572 GCACTGCCTGAAATTTCCCTC-----------------ACCCCCAACTTTCCCCACTGGC 2614

Qy   3037 TCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCC 3096
Db   2615 TCCACAACCCTGTTTGGAGCTGTTGCAGGACCAGAAGCACAAAGTGTGGTTTCCCAGGCC 2674

Qy   3097 TTTGTCCATCTCAG-CCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCA 3155
Db   2675 TTTGTCCATCTCAGCCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCA 2734

Qy   3156 GGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGT 3215
Db   2735 GGAGCACCCCCTGCCTGAGCTAA-GGAGGTCTTATCTCTCA-GGGGGGTTTAAGTGCCGT 2792

Qy   3216 TTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAA 3275
Db   2793 TTGCAATAAT---GTCTTATTTATTTAGCGGGGCAAATATTTTATACTGTAAGTGAGCAA 2849

Qy   3276 TCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAA 3335
Db   2850 TC--AGTATAATGTTTATGGTGATGAAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAA 2907

Qy   3336 AAAAAAAAAA 3345
Db   2908 AAAAAAAAAA 2917
```

```
RESULT 12
AX200994
LOCUS       AX200994                2904 bp    DNA     linear   PAT 29-AUG-2001
DEFINITION  Sequence 624 from Patent WO0151633.
ACCESSION   AX200994
VERSION     AX200994.1  GI:15390821
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2904)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Fanger,G.R., Day,C.H., Retter,M.W.,
            Stolk,J.A., Skeiky,Y.A., Wang,A. and Meagher,M.J.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0151633-A 624 19-JUL-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..2904
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      542 a    875 c    773 g    714 t
ORIGIN

  Query Match            64.4%;  Score 2196.4;  DB 6;  Length 2904;
  Best Local Similarity  88.2%;  Pred. No. 0;
  Matches 2560;  Conservative    0;  Mismatches    1;  Indels  343;  Gaps    2;

Qy    770 GTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC 829
Db      1 GTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC 60

Qy    830 TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTG 889
Db     61 TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTG 120

Qy    890 CTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCG 949
Db    121 CTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCG 180

Qy    950 CTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGT 1009
Db    181 CTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGT 240

Qy   1010 CCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAG 1069
Db    241 CCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAG 300

Qy   1070 CTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGG 1129
Db    301 CTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGG 360

Qy   1130 ATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAG 1189
```

```
Qy     338 AGAGCTGGGCTTTCCGGT---GCCGCAGCAGGCGGCTCA-----------------CCC 300
                  |||    |||  |||      ||||||||                  |||
Db     716 ProProGlyAlaGlnGlyProAlaGlyProGlyGlySerArgGlyAspProGlyProPro 735

Qy     299 ACAGCCTCTGGACCATAGTGGGCCAGGCGGG----------------------------- 269
              |||  |||||| :::||||||||||||
Db     736 GlyAlaGlnGlyPro-AlaGlyProGlyGlySerArgGlyAspProGlyProProGlyAl 755

Qy     268 ----------TAGGGCTCAGGGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCTCGGCTC 219
                       |||   ||||||     |||   |||
Db     755 aGlnGlyProAlaGlyProGlyGlySerArgGlyAspPro-------------------- 768

Qy     218 TGCTCCAGAAGCTGCGGCCTCTCCT------CCTTGCTGCCGCCAACTGCCTAGGAATCA 165
               |||     |||      |||      |||              |||
Db     769 -GlyProProGlyAlaHisGlyProAlaGlyProLysGlyAlaHisGlyProAlaGlyPr 788

Qy     164 GCCAGGCGCCCATTTC---TGCCAGCCCTTTGGTGCC---GGTCCAGCTTCT---CAGCC 114
               |||||||||           |||  ||||||   |||||||||     :::
Db     788 oLysGlyAlaHisGlyProAlaGlyProLysGlyAlaHisGlyProAlaGlyProLys-G 808

Qy     113 CATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTCATCACTCAGATC 54
                     ||||||      |||        ||||||                 |
Db     808 lyAlaGlnGlyProAlaGlyProGlyGlySerArgGlyAspPro---------GlyProP 825

Qy      53 CTGGCCGAGGCGCGCGGCTGTCACCCGGAGCCAGC 19
           |||||                 ||||||  |||
Db     825 roGlyAlaGlnGlyProAlaGlyProGlyGlySer 836
```

RESULT 42
US-08-931-820-4
; Sequence 4, Application US/08931820
; Patent No. 6010863
; GENERAL INFORMATION:
;   APPLICANT:
;   TITLE OF INVENTION: Assay for collagen degradation
;   NUMBER OF SEQUENCES: 4
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/931,820
;   FILING DATE:
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: EP 96202596.1
;   FILING DATE:
; INFORMATION FOR SEQ ID NO: 4:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1057 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   HYPOTHETICAL: NO
;   ORIGINAL SOURCE:
;   ORGANISM: Homo sapiens
;   TISSUE TYPE: Collagen type III
;   FEATURE:
;   NAME/KEY: Modified-site
;   LOCATION: 1055
;   OTHER INFORMATION: /label= Modified
;   OTHER INFORMATION: /note= "Ala may be Pro"
US-08-931-820-4

Alignment Scores:

Pred. No.:	6.85e-19	Length:	1057
Score:	351.50	Matches:	289
Percent Similarity:	29.85%	Conservative:	57
Best Local Similarity:	24.94%	Mismatches:	391
Query Match:	5.65%	Indels:	424
DB:	3	Gaps:	62

US-09-759-143-110 (1-3410) x US-08-931-820-4 (1-1057)

```
Qy    3051 AAACAGGGTTGTGGAGCTGGTGGGGAAAGTTGGGGGTAGGGGAAAGTTGGGGGTAGGGGA 2992
           |||   |||         ||||||    :::        |||      |||       |||
Db      13 LysSerGlyValAlaValGlyGlyLeuAlaGlyTyrProGlyProAlaGlyProProGly 32

Qy    2991 AATTTTGGGCAGTGCCTTCATCAGCCCAGTCCTAGAGAGAGTAGAGGGG----------- 2943
                                   |||    |||            |||
Db      33 --------------------ProProGlyPro-ProGlyThrSerGlyHisProGlySe 45

Qy    2942 -AGTGGAAGTGGGGGGAACCAGGCTGGGCCAAGAGAAGAGGGGTGGTTAGGGAAGCCGTT 2884
               ||||||  |||   |||       |||   |||   |||      |||   |||
Db      45 rProGlySerProGlyTyrGlnGlyProProGlyGluProGlyGlnAlaGly---ProSe 64

Qy    2883 GAGACCTGAAGCCCCACCCTCTACCTTCCTTCAACACCCTAACCTTGGGTAACAGCATTT 2824
               |||      ||||||
Db      64 rGlyProProGlyProProGly-------------------------------------- 71

Qy    2823 GGAATTATCATTTGGGATGAGTAGAATTTCCAAGGTCCTGGGTTAGGCATTTTGGGGGGC 2764
                  :::   |||   |||      :::         |||   :::
Db      72 -------AlaIleGlyProSerGlyProAlaGlyLysAspGlyGluSerGlyArgProGl 89

Qy    2763 CAGACCCCAGGAGAAGAAGATTCTGGCAATGATCAGCCCAATGACCAGCTATCTCAGGGG 2704
           ||||||    |||:::   :::               |||            ::::::||
Db      89 yArgProGlyGluArgGlyLeu--------------ProGlyProProGlyIleLysGl 104

Qy    2703 ACCTGATTGTTGGGGATCCCCCACCCTACCCAAATATTAGACACCAACACAGA------- 2651
           ||||        |||    |||    :::|||   :::::       ||||||
Db     104 yProAla------GlyIle-ProGlyPheProGlyMetLys---GlyHisArgGlyPheA 121

Qy    2650 ----------------AAAGCTAGCAATGGATTCCCTTCTACTTTGTTAAATAAATAAG 2608
                            |||         |||   |||                     :
Db     121 spGlyArgAsnGlyGluLysGlyGluThrGlyAlaProGly-----------------L 135

Qy    2607 TTAAATATTTAAATGCCTGTGTCTCTGTGATGGCAACAGAAGGACCAACAGGCCACATCC 2548
           ::|||      |||                           ||||||  |||
Db     135 euLysGlyGluAsnGlyLeuProGlyGluAsnGlyAlaProGlyProMetGlyProArgG 155

Qy    2547 TGATAAAAGGTAAGAGGGGGGTGGATCAGCAAAAAGACAGTGCTGTGGGCTGAGGGGACC 2488
               |||:::||||||                      :::|||         |||:::
Db     155 lyAlaProGlyGluArgGlyArgProGlyLeuProGlyAlaAlaGlyAlaArgGlyAsn- 174

Qy    2487 TGGTTCTTGTGTGTTGCCCCTCAGGACTCTTCCCCTACAAATAAGTCATATGTTCAAATC 2428

Db     174 ------------------------------------------------------------ 174

Qy    2427 CCATGGAGGAGTGTTTCATCCTAGAAACTCCCATGCAAGAGCTACATTAAACGAAGCTGC 2368

Db     174 ------------------------------------------------------------ 174

Qy    2367 AGGTTAAGGGGCTTAGA------GATGGGAAACCAGGTGACTGAGTTTATTCAGCTCCCA 2314
                   |||   |||      ||||||:::||||||                  |||
Db     175 -----AspGlyAlaArgGlySerAspGlyGlnProGly------------------ProP 187

Qy    2313 AAAACCCTTCTCTAGGTGTGTCTCAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTA 2254
                |||      |||   :::                                  |||:
Db     187 roGlyProPro---GlyThrAlaGlyPhePro------------------------GlyS 198

Qy    2253 ATCCACCTGCAGAGTCCCCGCATTCCAGTGCATGGAGCCCT------TCTGGCCTCCCTG 2200
           ::|||   |||:::              |||   ||||||       :::|||  |||
Db     198 erProGlyAlaLysGlyGluValGlyProAlaGlySerProGlySerAsnGlyAlaProG 218

Qy    2199 TATAAGTCCAGACTGAAACCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGG 2140
                                              |||  |||    |||         |
Db     218 ly--------------------------------GlnArgGlyGluProGlyProGlnG 227

Qy    2139 GAGAG------AGGAGAGGGACGCCCCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAG 2086
           ||            :::|||   |||   ||||||   :::           :::
Db     227 lyHisAlaGlyAlaGlnGlyProProGlyProProGlyIle-----------AsnGlyS 243

Qy    2085 CACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCC 2026
           ::   ||||||      |||                                ||||||
```

```
Db    114 ---GlyGly--------------SerArgGlyAspProGlyProProGlyAlaGlnGly 127
Qy   2248 CCTGCAGAGTCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAG 2189
Db    128 ProAla---GlyProGlyGlySerArgGlyAspPro---GlyProPro----------- 141
Qy   2188 ACTGAAACCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGA 2129
Db    142 -----------------GlyAlaGln------GlyProAlaGlyProGlyGlySerArg 153
Qy   2128 GGGACGCCCCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGTGGCAGCAGA 2069
Db    154 GlyAspProGlyProProGlyAlaGln------GlyProAlaGlyProGlyGlySerArg 171
Qy   2068 ---GAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGC-------- 2020
Db    172 GlyAspProGlyProProGly-AlaGlnGlyProAlaGlyProGlyGlySerArgGlyAs 191
Qy   2019 -------CCCATGGGGCTAACAGGA---GCGGGGAGCTGGGACCCAGTGAGGCAGGCCCT 1970
Db    191 pProGlyProProGlyAlaGlnGlyProAla-GlyProGlyGlySerArgGlyAspProG 211
Qy   1969 CCACCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAA 1910
Db    211 lyProProGlyAlaGlnGly---------------------------------------- 217
Qy   1909 TACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCCTGCGGCAGACACCATATA 1850
Db    218 -----ProAlaGlyProGlyGlySerArgGlyAspProGly---------ProProGlyA 233
Qy   1849 GGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCATAAACAGGGATGGGGCCACCTG 1790
Db    233 laGln---------GlyProAlaGlyProGlyGlySerArgGlyAspProGlyProProG 250
Qy   1789 GGACAGCAGGAAGGCACTATCCAGGATGGCGA-----------GGTCCAGGCAGATGCC 1742
Db    250 lyAlaGlnGlyProAlaGlyProGlyGlySerArgGlyAspProGlyProProGlyAlaH 270
Qy   1741 CCGGCCCGGAA--------------------CCACCCTGGCCTCGGTGGGCTCACCCAC 1703
Db    270 isGlyProAlaGlyProLysGlyAlaHisGlyProAlaGlyProLysGlyAlaHisGlyP 290
Qy   1702 CACCACACGTACGGAGACATCACAG------------------GCAGAGGCCCCGCAGA 1662
Db    290 roAlaGlyProLysGlyAlaHisGlyProAlaGlyProLysGlyAlaGlnGlyPro---- 308
Qy   1661 GCGCGGGTGGAGGTGGGAGCAGG------CCACTGCCTCCAGCACCCACGTGTCCATTAG 1608
Db    309 --AlaGlyProGlyGlySerArgGlyAspProGlyProProGlyAlaGlnGlyProAlaG 328
Qy   1607 GG-----------AAGGGAGCTCCAGGC--------------------TTAGGGCCTG 1581
Db    328 lyProGlyGlySerArgGlyAspProGlyProProGlyAlaGlnGlyProAlaGlyProG 348
Qy   1580 GCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGTATT 1521
Db    348 ly-Gly------SerArgGlyAspProGlyProProGlyAlaGlnGlyProAlaGlyPro 365
Qy   1520 TGGGCA------GGAACACCTGCTTCTCCCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCA 1467
Db    366 GlyGlySerArgGlyAspProGlyProProGlyAlaGlnGly-----------ProAla 381
Qy   1466 GGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCA 1407
Db    382 GlyProGlyGlySerArg-GlyAspProGlyProProGly------------------- 394
Qy   1406 CACTGTGGGACAGGCATGTGGCACCGGCAGCCACAGGGAAAGCTGCCACACTGGCCAAAT 1347
Db    395 -AlaGlnGlyProAla---GlyProGlyGlySerArgGlyAsp--ProGlyProProGly 412
Qy   1346 AGACTGCTCGAGTGCCGAATCGCTGCACCAGCCGGTCCATGACCAGAGAGA--------- 1296
Db    413 -----------------AlaGlnGlyProAlaGlyProGlyGlySerArgGlyAspPro 426

Qy   1295 ---AGACCAGGGAGATGGCGCACTGCAGGAACAGCC--------------CCAGGCTGC 1254
Db    427 GlyProProGlyAlaGlnGlyProAlaGlyProGlyGlySerArgGlyAspProGlyPro 446
Qy   1253 CCATCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCA 1194
Db    447 Pro-----------------GlyAlaGlnGlyProAlaGlyProGlyGlySerArgGly 460
Qy   1193 CGCCCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTG 1134
Db    461 AspProGly---ProProGlyAlaGlnGlyProAlaGly-------------------- 472
Qy   1133 CCATCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGC 1074
Db    473 ProGlyGlySerArgGlyAspProGlyProProGlyAlaGln---GlyProAlaGly--- 490
Qy   1073 ACAGCTGGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGGCCCGGC 1014
Db    491 ProGlyGlySerArgGlyAspProGlyProProGlyAlaGlnGlyProAlaGlyProGly 510
Qy   1013 ---------ATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCT 963
Db    511 GlySerArgGlyAspProGlyProProGlyAlaHisGlyProAlaGlyProLysGlyAla 530
Qy    962 CGGTGGGGCCCAGCGCTGCCT--------------CCTCAGCCACCAGCAGTGTGGCTG 918
Db    531 His---GlyProAlaGlyProLysGlyAlaHisGlyProAlaGlyProLys--------- 546
Qy    917 CTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGCACT-------------- 873
Db    547 -----------GlyAlaHisGlyProAlaGlyProLysGlyAlaGlnGlyProAlaGly 562
Qy    872 --------------------------CCTCCTGGGTGCCCAGGT---AGGGGGCCAGGG 843
Db    563 ProGlyGlySerArgGlyAspProGlyProProGlyAlaGlnGlyProAlaGlyPro--- 581
Qy    842 CACTGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGACTGATCA 783
Db    582 --------------------------GlyGlySerArgGlyAspPro------------ 588
Qy    782 TGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGTCCG---GGTCCCGGAAGAGGTCAG 726
Db    589 --------------GlyProProGlyAlaGlnGlyProAlaGlyProGly---GlySer 602
Qy    725 AGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCA--- 669
Db    603 ArgGlyAspPro-----------GlyProProGlyAlaGlnGlyProAlaGlyProGly 618
Qy    668 --------------------CGCCCAGGATGAGCAGTGCCAGCT---CCAGGGGCC--- 636
Db    619 GlySerArgGlyAspProGlyProProGlyAlaGlnGlyProAlaGlyProGlyGlySer 638
Qy    635 ---TGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGC 579
Db    639 ArgGlyAspProGlyProProGlyAlaGlnGlyProAlaGlyProGlyGlySerArgGly 658
Qy    578 TCAGCAGGATGCCCAAGGACAGTGCCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCAC 519
Db    659 AspProGlyProPro-------------------------GlyAlaGlnGlyProAla 669
Qy    518 GCCAGTGGTCACTGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAA 459
Db    670 GlyProGly--------------GlySerArgGlyAspProGlyProProGlyAlaGln 684
Qy    458 TGCCCAGCACCATGGTCATGAACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCA 399
Db    685 GlyProAlaGlyProGlyGlySerArgGlyAspProGlyProProGlyAlaGlnGly--- 703
Qy    398 CATAGGTGATGCCTGCGGCCAAACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCA 339
Db    704 -----------------ProAlaGlyProGlyGlySerArg------GlyAspProGly 715
```

```
                 |||       |||     |||             ||||||:::
Db     360 ProGlySerProGlyAlaProGlyThrProGlyPro---GlnGlyLeuProGlySerPro 378

Qy    1178 CCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGC------ 1125
                 |||       |||                  :::|||   |||
Db     379 GlyAlaProGlyThrProGlyProGlnGly-----------LeuProGlySerProGly 394

Qy    1124 ------TGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACA 1071
                 |||     |||:::               ||||||   |||:::
Db     395 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 414

Qy    1070 GCT------GGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGG--- 1020
                 |||      |||     |||   ||||||:::      |||   |||
Db     415 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 434

Qy    1019 CCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGG 960
           ||||||          :::     |||      ||||||    |||        :::
Db     435 ProGlySerProGlyAlaProGlyThr---ProGlyProGlnGlyLeuProGlySerPro 453

Qy     959 TGGGGCCCAGCGCTGCCT---CCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGA 903
                 |||     |||  ||||||                               |||
Db     454 GlyAlaProGlyThrProGlyProGln-----------------------------Gly 463

Qy     902 AGATGAGGGTGAGCAGGCCAAAGAGGCACTCCT--------------CCTGGGTGCCCA 858
                 |||     |||     |||||||||                ||||||   |||
Db     464 LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro 483

Qy     857 GGTAGGGGGCCAGGGCACTGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGC 798
           |||      ||||||          |||                  ||||||||||||:::
Db     484 Gly---AlaProGly------ThrProGlyProGlnGlyLeuProGlySerProGlyAla 500

Qy     797 CCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGT 741
           |||                                 ||||||    |||    ||||||
Db     501 Pro-------------------------GlyThrProGlyProGlnGlyLeuProGly 511

Qy     740 ---CCCGGAAGAGGTCAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGA 684
              ||||||          ||||||                  :::||||||
Db     512 SerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGly------ 529

Qy     683 AGTCCAGCAGCCCCACGCCCAGGATGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGC 624
           :::|||      |||   |||               :::|||||||||       ||||||
Db     530 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla------ProGly 547

Qy     623 ACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCA 564
           |||                         |||    |||        ||||||   |||   |||
Db     548 ThrProGly---------------ProGlnGly---LeuProGlySerProGlyAlaPro 561

Qy     563 AGGACAGTG---CCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCAC 507
                |||     |||:::          |||
Db     562 GlyThrProGlyProGlnGlyLeuProGly----------------------------- 571

Qy     506 TGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCA 447
                 |||   ||||||         |||||||||                      |||
Db     572 ------SerProGlyAlaProGlyThrProGlyProGln-----------GlyLeuPro 585

Qy     446 TGGTCATGAACTTCTCCTCTACCCCCA-----------CTTCCAGCAGCAGAGGCGGCA 399
              |||      :::|||     |||          ||||||   :::        |||
Db     586 GlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 605

Qy     398 CATAGGTGATGCCTGCGGCCAAACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCA 339
                                   |||  ||||||     |||   |||   |||   |||
Db     606 -----------------ProGlyThrProGlyProGlnGlyLeuProGlySerProGly 619

Qy     338 AGAGCTGGGCTTTCCGGT---GCCGCAGCAGGCGGCTCACCCACAGCCTCTGGACCATAG 282
              |||      |||               |||||||||   |||    |||
Db     620 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr-Pr 639

Qy     281 TGGGCCAGGCGGG---TAGGGCTCAGGGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCT 225
             ||||||   |||        ||||||   |||      |||||||||          ||
Db     639 oGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLe 659

Qy     224 CGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAACTGCCTAGGAATCA 165
            |        :::|||   |||        :::|||                  |||:::
```

```
Db     659 uProGlySerProGlyAlaProGlyThrProGly--------------ProGlnGlyLe 674

Qy     164 GCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGCTTCTCAGCCCATGCTCA- 106
              ||||||:::            |||         ||||||          |||   |||
Db     674 uProGlySerProGlyAla---ProGlyThrProGlyProGlnGly-LeuProGlySerP 693

Qy     105 -------------ACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTCATCACT 60
                        ||||||           |||    |||    |||
Db     693 roGlyAlaProGlyThrProGlyProGlnGly---LeuProGlySerProGlyAlaProG 712

Qy      59 CAGATCCTGGC---CGAGGCGCGCGGCTGTCACCCGGAGCC 22
                 ||||||   :::|||        |||||||||||
Db     712 lyThrProGlyProGlnGlyLeuProGlySerProGlyAla 725
```

```
RESULT 41
US-08-642-255-72
; Sequence 72, Application US/08642255
; Patent No. 5773249
;  GENERAL INFORMATION:
;    APPLICANT:  CAPPELLO, Joseph
;    APPLICANT:  FERRARI, Franco A.
;    TITLE OF INVENTION:  High Molecular Weight Collagen-Like
;    TITLE OF INVENTION:  Protein Polymers
;    NUMBER OF SEQUENCES:  135
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/642,255
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROWLAND, Bertram I.
;      REGISTRATION NUMBER:  20,015
;      REFERENCE/DOCKET NUMBER:  A55556-3/BIR
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 494-8700
;      TELEFAX:  (415) 494-8771
;      TELEX:  910 277299 FHT UR
;  INFORMATION FOR SEQ ID NO:  72:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1065 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;  MOLECULE TYPE:  protein
US-08-642-255-72

Alignment Scores:
Pred. No.:               4.82e-19      Length:         1065
Score:                   353.50        Matches:        270
Percent Similarity:      32.33%        Conservative:   20
Best Local Similarity:   30.10%        Mismatches:     332
Query Match:             5.68%         Indels:         276
DB:                      1             Gaps:           61

US-09-759-143-110 (1-3410) x US-08-642-255-72 (1-1065)

Qy    2347 GGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCTCAA 2288
              |||   ||||||                        |||       |||
Db     100 GlyAspProGlyProProGlyAlaGlnGlyProAlaGlyPro------------------ 113

Qy    2287 CTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCA-------------------- 2249
              ||||||                  |||   |||:::|||
```

```
Qy        164 GCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGCTTCTCAGCCCATGCTCA- 106
                 ||||||:::          |||       ||||||          |||   |||
Db        674 uProGlySerProGlyAla---ProGlyThrProGlyProGlnGly-LeuProGlySerP 693

Qy        105 -------------ACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTCATCACT 60
                           ||||||         |||    |||    |||
Db        693 roGlyAlaProGlyThrProGlyProGlnGly---LeuProGlySerProGlyAlaProG 712

Qy         59 CAGATCCTGGC---CGAGGCGCGCGGCTGTCACCCGGAGCC 22
                      ||||||    :::|||         |||||||||||
Db        712 lyThrProGlyProGlnGlyLeuProGlySerProGlyAla 725
```

RESULT 40
US-08-397-633A-50
; Sequence 50, Application US/08397633A
; Patent No. 5773577
; GENERAL INFORMATION:
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: PRODUCTS COMPRISING SUBSTRATESCAPABLE
; TITLE OF INVENTION: OF ENZYMATIC CROSS-LINKING
; NUMBER OF SEQUENCES: 105
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/397,633A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A-58848-1/BIR PROP-011-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 897 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-397-633A-50

Alignment Scores:

Pred. No.:	2.67e-19	Length:	897
Score:	356.50	Matches:	263
Percent Similarity:	35.16%	Conservative:	32
Best Local Similarity:	31.35%	Mismatches:	327
Query Match:	5.73%	Indels:	218
DB:	1	Gaps:	57

US-09-759-143-110 (1-3410) x US-08-397-633A-50 (1-897)

```
Qy       2347 GGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCTCAA 2288
                |||  ||||||                    ||||||       |||     |||
Db         46 GlySerProGly---------------AlaProGlyThrPro------GlyProGlnGly 58

Qy       2287 CTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCAGAGTCCCCGCATTCC 2228
                |||   |||                  |||   |||   |||             |||
Db         59 LeuProGly----------------SerProGlyAlaProGlyThrProGlyProGlnGly 73

Qy       2227 AGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGC 2168
                          ||||||   |||   |||              |||       |||   :::|||
Db         74 LeuProGlySerPro---GlyAlaPro---------GlyThrProGlyPro---GlnGly 88

Qy       2167 CTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCCCCAGCCCCCAGCT 2108
              |||       |||||||||                          |||   |||      |||
Db         89 Leu---ProGlySerPro--------------------GlyAlaProGlyThrProGly 100

Qy       2107 GTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGC 2048
                 |||                |||                |||             |||
Db        101 ProGln------------GlyLeuProGlySerProGlyAlaProGlyThrProGly--P 116

Qy       2047 AACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGGA 1988
                |||   |||     .||||||   ||||||                   :::|||
Db        116 roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 136

Qy       1987 CCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGC 1928
                      |||   |||   ||||||
Db        136 lySerProGlyAlaProGlyThrPro---------------------------------- 144

Qy       1927 CAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCC 1868
                           |||   ||||||                            ||||||
Db        145 --------GlyProGlnGlyLeuPro-------------------GlySerProGlyA 155

Qy       1867 TGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCATAAA 1808
                  ||||||   |||  .      |||                      ||||||
Db        155 laProGlyThrProGlyProGlnGlyLeuProGly---------------SerPro---- 168

Qy       1807 CAGGGATGGGGCCACCTGGGACAGCAGGAAGGCAC---------------------TATC 1769
                      |||   |||||||||   |||                               |
Db        169 --------GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaP 186

Qy       1768 CAGGATGGCGAGGTCCAGGCAGATGCCCCGGC------------CCGGAACCACCCTGGC 1721
              |||||      ||||||           ||||||          |||      |||   |
Db        186 roGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyP 206

Qy       1720 CTCGGTG-------GGCTCACCCACCACCACACGTACGGAGACATCACAGGCA------G 1674
              ||:::          |||||||||          |||        |||
Db        206 roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 226

Qy       1673 AGGCCCCGCAGAGCGCGGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTC 1614
                :::|||   :::   |||       |||   :::   ||||||   :::|||         |
Db        226 lySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaP 246

Qy       1613 CATTAGGGAAGGGAGCTCCAGGCTTAGGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCT 1554
              ||        |||      ||||||   ||||||          :::   |||   |||
Db        246 roGlyThrProGlyProGlnGlyLeu---ProGly-SerProGlyAlaProGlyThrPro 264

Qy       1553 CACTGCTAGCACCTCCAGTGTCCCCTCGGTATT------TGGGCAGGAACACCTGCTTCT 1500
                                        |||   |||:::          |||   |||   :::
Db        265 Gly-----------------ProGlnGlyLeuProGlySerProGlyAlaProGlyThr 278

Qy       1499 CCCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACC 1440
              ||||||   :::|||   |||             ||||||   |||         |||   ||
Db        279 ProGlyProGlnGlyLeuPro------------GlySerProGlyAla-ProGlyThrPr 294

Qy       1439 CGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACC-- 1382
              ||||      |||               |||                 |||   ||||||
Db        294 oGlyProGlnGlyLeuProGlySerProGly---------------AlaProGlyThrPr 309

Qy       1381 -GGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCT 1323
                 |||  |||||||||   ||||||      |||
Db        309 oGly--ProGlnGly---LeuProGlySerPro-----------------------Gly 319

Qy       1322 GCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACA--- 1266
              ||||||      |||   |||:::      ||||||   |||            ||||||
Db        320 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 339

Qy       1265 ----------------------------GCCCCAGGCTGCCCATCC---GAACGCCTT 1239
                                          |||||||||   |||         :::   |||
Db        340 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 359

Qy       1238 CATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTACAGCC 1179
```

```
                    ||||||   |||   |||              |||       |||     ::::|||
Db      74 LeuProGlySerPro---GlyAlaPro---------GlyThrProGlyPro---GlnGly 88

Qy    2167 CTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCCCCAGCCCCCAGCT 2108
           |||      |||||||||                           |||   |||      |||
Db      89 Leu---ProGlySerPro--------------------GlyAlaProGlyThrProGly 100

Qy    2107 GTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGC 2048
             |||                     |||          |||         |||
Db     101 ProGln------------GlyLeuProGlySerProGlyAlaProGlyThrProGly--P 116

Qy    2047 AACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGGA 1988
              |||   |||     ||||||   ||||||                :::|||
Db     116 roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 136

Qy    1987 CCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGC 1928
                   |||   |||   ||||||
Db     136 lySerProGlyAlaProGlyThrPro------------------------------------ 144

Qy    1927 CAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCC 1868
                   |||   ||||||                               ||||||
Db     145 --------GlyProGlnGlyLeuPro--------------------GlySerProGlyA 155

Qy    1867 TGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCATAAA 1808
                 ||||||    |||         |||                  ||||||
Db     155 laProGlyThrProGlyProGlnGlyLeuProGly---------------SerPro---- 168

Qy    1807 CAGGGATGGGGCCACCTGGGACAGCAGGAAGGCAC--------------------TATC 1769
                   |||   |||||||||   |||                              |
Db     169 --------GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaP 186

Qy    1768 CAGGATGGCGAGGTCCAGGCAGATGCCCCGGC------------CCGGAACCACCCTGGC 1721
           |||||     ||||||          ||||||            |||     |||   |
Db     186 roGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyP 206

Qy    1720 CTCGGTG-------GGCTCACCCACCACCACACGTACGGAGACATCACAGGCA------G 1674
           ||:::           |||||||||              |||        |||
Db     206 roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 226

Qy    1673 AGGCCCCGCAGAGCGCGGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTC 1614
              :::|||   :::   |||       |||   :::   ||||||   :::|||      |
Db     226 lySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaP 246

Qy    1613 CATTAGGGAAGGGAGCTCCAGGCTTAGGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCT 1554
           ||          |||       ||||||  ||||||          :::   |||   |||
Db     246 roGlyThrProGlyProGlnGlyLeu---ProGly-SerProGlyAlaProGlyThrPro 264

Qy    1553 CACTGCTAGCACCTCCAGTGTCCCCTCGGTATT------TGGGCAGGAACACCTGCTTCT 1500
                                 |||   |||:::            |||   |||   :::
Db     265 Gly------------------ProGlnGlyLeuProGlySerProGlyAlaProGlyThr 278

Qy    1499 CCCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACC 1440
           ||||||   :::|||   |||              ||||||   |||      |||   ||
Db     279 ProGlyProGlnGlyLeuPro------------GlySerProGlyAla-ProGlyThrPr 294

Qy    1439 CGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACC-- 1382
           ||||     |||                |||               |||   ||||||
Db     294 oGlyProGlnGlyLeuProGlySerProGly---------------AlaProGlyThrPr 309

Qy    1381 -GGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCT 1323
              |||  |||||||||   ||||||      |||
Db     309 oGly--ProGlnGly---LeuProGlySerPro-----------------------Gly 319

Qy    1322 GCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACA--- 1266
           ||||||       |||   |||:::         ||||||                ||||||
Db     320 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 339

Qy    1265 -----------------------------GCCCCAGGCTGCCCATCC---GAACGCCTT 1239
                                        |||||||||   |||      :::   |||
Db     340 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 359

Qy    1238 CATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTACAGCC 1179
                   |||          |||       |||                  ||||||:::

Db     360 ProGlySerProGlyAlaProGlyThrProGlyPro---GlnGlyLeuProGlySerPro 378

Qy    1178 CCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGC------ 1125
                |||      |||                               :::|||   |||
Db     379 GlyAlaProGlyThrProGlyProGlnGly------------LeuProGlySerProGly 394

Qy    1124 ------TGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACA 1071
                    |||      |||:::                  ||||||    |||:::
Db     395 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 414

Qy    1070 GCT------GGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGG--- 1020
                    |||     |||      |||     ||||||:::       |||   |||
Db     415 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 434

Qy    1019 CCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGG 960
           ||||||        :::        |||        ||||||      |||        :::
Db     435 ProGlySerProGlyAlaProGlyThr---ProGlyProGlnGlyLeuProGlySerPro 453

Qy     959 TGGGGCCCAGCGCTGCCT---CCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGA 903
              |||      |||    ||||||                               |||
Db     454 GlyAlaProGlyThrProGlyProGln-----------------------------Gly 463

Qy     902 AGATGAGGGTGAGCAGGCCAAAGAGGCACTCCT---------------CCTGGGTGCCCA 858
              |||      |||      |||||||||                   ||||||   |||
Db     464 LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro 483

Qy     857 GGTAGGGGGCCAGGGCACTGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGC 798
           |||      ||||||         |||                ||||||||||||:::
Db     484 Gly---AlaProGly------ThrProGlyProGlnGlyLeuProGlySerProGlyAla 500

Qy     797 CCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGT 741
           |||                               ||||||      |||   ||||||
Db     501 Pro------------------------GlyThrProGlyProGlnGlyLeuProGly 511

Qy     740 ---CCCGGAAGAGGTCAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGA 684
              ||||||               ||||||               :::||||||
Db     512 SerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGly------ 529

Qy     683 AGTCCAGCAGCCCCACGCCCAGGATGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGC 624
            :::|||      |||   |||               :::|||||||||       ||||||
Db     530 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla------ProGly 547

Qy     623 ACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCA 564
           |||                     |||   |||         ||||||   |||   |||
Db     548 ThrProGly---------------ProGlnGly---LeuProGlySerProGlyAlaPro 561

Qy     563 AGGACAGTG---CCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCAC 507
              |||      |||:::          |||
Db     562 GlyThrProGlyProGlnGlyLeuProGly------------------------------ 571

Qy     506 TGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCA 447
                 |||   ||||||            |||||||||                  |||
Db     572 ------SerProGlyAlaProGlyThrProGlyProGln-----------GlyLeuPro 585

Qy     446 TGGTCATGAACTTCTCCTCTACCCCCA------------CTTCCAGCAGCAGAGGCGGCA 399
              |||      :::|||   |||                   ||||||   :::      |||
Db     586 GlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 605

Qy     398 CATAGGTGATGCCTGCGGCCAAACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCA 339
                                |||   ||||||   |||   |||   |||   |||
Db     606 ------------------ProGlyThrProGlyProGlnGlyLeuProGlySerProGly 619

Qy     338 AGAGCTGGGCTTTCCGGT---GCCGCAGCAGGCGGCTCACCCACAGCCTCTGGACCATAG 282
                 |||      |||                |||||||||   |||   |||
Db     620 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr-Pr 639

Qy     281 TGGGCCAGGCGGG---TAGGGCTCAGGGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCT 225
              ||||||   |||      ||||||   |||      |||||||||            ||
Db     639 oGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLe 659

Qy     224 CGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAACTGCCTAGGAATCA 165
           |         :::|||   |||      :::|||                     |||:::
Db     659 uProGlySerProGlyAlaProGlyThrProGly---------------ProGlnGlyLe 674
```

```
Qy        959 TGGGGCCCAGCGCTGCCT---CCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGA 903
                   |||     |||   ||||||                             |||
Db        454 GlyAlaProGlyThrProGlyProGln-----------------------------Gly 463

Qy        902 AGATGAGGGTGAGCAGGCCAAAGAGGCACTCCT---------------CCTGGGTGCCCA 858
                   |||     |||    |||||||||                ||||||   |||
Db        464 LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro 483

Qy        857 GGTAGGGGGCCAGGGCACTGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGC 798
               |||     ||||||            |||                ||||||||||||:::
Db        484 Gly---AlaProGly------ThrProGlyProGlnGlyLeuProGlySerProGlyAla 500

Qy        797 CCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGT 741
               |||                                  ||||||    |||   ||||||
Db        501 Pro--------------------------GlyThrProGlyProGlnGlyLeuProGly 511

Qy        740 ---CCCGGAAGAGGTCAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGA 684
                 ||||||            ||||||                  :::||||||
Db        512 SerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGly------ 529

Qy        683 AGTCCAGCAGCCCCACGCCCAGGATGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGC 624
               :::|||     |||   |||             :::||||||||||      ||||||
Db        530 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla------ProGly 547

Qy        623 ACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCA 564
               |||                        |||   |||        ||||||  |||   |||
Db        548 ThrProGly--------------ProGlnGly---LeuProGlySerProGlyAlaPro 561

Qy        563 AGGACAGTG---CCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCAC 507
                 |||     |||:::          |||
Db        562 GlyThrProGlyProGlnGlyLeuProGly----------------------------- 571

Qy        506 TGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCA 447
                    |||   ||||||          ||||||||               |||
Db        572 ------SerProGlyAlaProGlyThrProGlyProGln-----------GlyLeuPro 585

Qy        446 TGGTCATGAACTTCTCCTCTACCCCCA-----------CTTCCAGCAGCAGAGGCGGCA 399
               |||      :::|||    |||          ||||||    :::        |||
Db        586 GlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 605

Qy        398 CATAGGTGATGCCTGCGGCCAAACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCA 339
                                   |||   ||||||    |||    |||    |||    |||
Db        606 -----------------ProGlyThrProGlyProGlnGlyLeuProGlySerProGly 619

Qy        338 AGAGCTGGGCTTTCCGGT---GCCGCAGCAGGCGGCTCACCCACAGCCTCTGGACCATAG 282
                    |||     |||             |||||||||    |||    |||
Db        620 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr-Pr 639

Qy        281 TGGGCCAGGCGGG---TAGGGCTCAGGGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCT 225
               ||||||  |||        ||||||   |||      |||||||||             ||
Db        639 oGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLe 659

Qy        224 CGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAACTGCCTAGGAATCA 165
               |     :::|||   |||       :::|||                |||:::
Db        659 uProGlySerProGlyAlaProGlyThrProGly--------------ProGlnGlyLe 674

Qy        164 GCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGCTTCTCAGCCCATGCTCA- 106
               ||||||:::            |||       ||||||         |||    |||
Db        674 uProGlySerProGlyAla---ProGlyThrProGlyProGlnGly-LeuProGlySerP 693

Qy        105 --------------ACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTCATCACT 60
                            ||||||        |||    |||    |||
Db        693 roGlyAlaProGlyThrProGlyProGlnGly---LeuProGlySerProGlyAlaProG 712

Qy         59 CAGATCCTGGC---CGAGGCGCGCGGCTGTCACCCGGAGCC 22
                   ||||||     :::|||          ||||||||||||
Db        712 lyThrProGlyProGlnGlyLeuProGlySerProGlyAla 725
```

RESULT 39
US-09-444-791A-103
; Sequence 103, Application US/09444791A
; Patent No. 6355776
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; Richardson, Charles
; Chambers, James
; Causey, Stuart
; Pollock, Thomas J.
; Cappello, Joseph
; Crissman, John W.
; TITLE OF INVENTION: No. 6355776el Peptides Comprising Repetitive
; Units of Amino Acids and DNA Sequences Encoding the Same
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr Hohbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/444,791A
; FILING DATE: 22-No. 6355776-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/482,085
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/175,155
; FILING DATE: 29-DEC-1993
; APPLICATION NUMBER: US 08/053,049
; FILING DATE: 22-APR-1993
; APPLICATION NUMBER: US 07/114,618
; FILING DATE: 29-OCT-1987
; APPLICATION NUMBER: US 06/927,258
; FILING DATE: 04-NOV-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-55186-11/RFT/BTC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 837 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 103:
US-09-444-791A-103

Alignment Scores:

Pred. No.:	2.61e-19	Length:	837
Score:	356.50	Matches:	263
Percent Similarity:	35.16%	Conservative:	32
Best Local Similarity:	31.35%	Mismatches:	327
Query Match:	5.73%	Indels:	218
DB:	4	Gaps:	57

US-09-759-143-110 (1-3410) x US-09-444-791A-103 (1-837)

```
Qy       2347 GGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCTCAA 2288
                 |||  ||||||                ||||||      |||         |||
Db         46 GlySerProGly--------------AlaProGlyThrPro------GlyProGlnGly 58

Qy       2287 CTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCAGAGTCCCCGCATTCC 2228
                 |||   |||                |||    |||    |||            |||
Db         59 LeuProGly--------------SerProGlyAlaProGlyThrProGlyProGlnGly 73

Qy       2227 AGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGC 2168
```

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/482,085B
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 06/927,258
;       FILING DATE:  04-NOV-1986
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/114,618
;       FILING DATE:  29-OCT-1987
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/053,049
;       FILING DATE:  22-APR-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/175,155
;       FILING DATE:  29-DEC-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Trecartin, Richard F.
;       REGISTRATION NUMBER:  31,801
;       REFERENCE/DOCKET NUMBER:  A-55186-6/RFT/MTK
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-781-1989
;       TELEFAX:  415-398-3249
;   INFORMATION FOR SEQ ID NO:  103:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  837 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-482-085B-103

Alignment Scores:
Pred. No.:              2.61e-19      Length:         837
Score:                  356.50        Matches:        263
Percent Similarity:     35.16%        Conservative:   32
Best Local Similarity:  31.35%        Mismatches:     327
Query Match:            5.73%         Indels:         218
DB:                     3             Gaps:           57

US-09-759-143-110 (1-3410) x US-08-482-085B-103 (1-837)

Qy       2347 GGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCTCAA 2288
               |||  ||||||                |||||     |||    |||
Db         46 GlySerProGly--------------AlaProGlyThrPro------GlyProGlnGly 58

Qy       2287 CTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCAGAGTCCCCGCATTCC 2228
               |||  |||                 |||  |||  |||          |||
Db         59 LeuProGly---------------SerProGlyAlaProGlyThrProGlyProGlnGly 73

Qy       2227 AGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGC 2168
                       ||||||  |||  |||          |||    |||  :::|||
Db         74 LeuProGlySerPro---GlyAlaPro---------GlyThrProGlyPro---GlnGly 88

Qy       2167 CTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCCCCAGCCCCCAGCT 2108
               |||      |||||||||                    |||   |||     |||
Db         89 Leu---ProGlySerPro--------------------GlyAlaProGlyThrProGly 100

Qy       2107 GTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGC 2048
                 |||                    |||          |||         |||
Db        101 ProGln------------GlyLeuProGlySerProGlyAlaProGlyThrProGly--P 116

Qy       2047 AACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGGA 1988
                 |||   |||     ||||||  ||||||           :::|||
Db        116 roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 136

Qy       1987 CCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGC 1928
                       |||   |||  ||||||
Db        136 lySerProGlyAlaProGlyThrPro---------------------------------- 144

Qy       1927 CAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCC 1868
                            |||  ||||||                        ||||||
Db        145 --------GlyProGlnGlyLeuPro--------------------GlySerProGlyA 155

Qy       1867 TGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCATAAA 1808
                 ||||||    |||        |||                ||||||
Db        155 laProGlyThrProGlyProGlnGlyLeuProGly--------------SerPro---- 168

Qy       1807 CAGGGATGGGGCCACCTGGGACAGCAGGAAGGCAC--------------------TATC 1769
                 |||  |||||||||  |||                                  |
Db        169 --------GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaP 186

Qy       1768 CAGGATGGCGAGGTCCAGGCAGATGCCCCGGC-----------CCGGAACCACCCTGGC 1721
              |||||    ||||||          ||||||            |||      |||   |
Db        186 roGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyP 206

Qy       1720 CTCGGTG-------GGCTCACCCACCACCACACGTACGGAGACATCACAGGCA------G 1674
              ||:::        |||||||||       |||         |||
Db        206 roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 226

Qy       1673 AGGCCCCGCAGAGCGCGGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTC 1614
                :::|||   :::    |||      |||    :::   ||||||   :::|||      |
Db        226 lySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaP 246

Qy       1613 CATTAGGGAAGGGAGCTCCAGGCTTAGGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCT 1554
              ||          |||     ||||||  ||||||           :::   |||  |||
Db        246 roGlyThrProGlyProGlnGlyLeu---ProGly-SerProGlyAlaProGlyThrPro 264

Qy       1553 CACTGCTAGCACCTCCAGTGTCCCCTCGGTATT------TGGCAGGAACACCTGCTTCT 1500
                                       |||  |||:::         |||   |||   :::
Db        265 Gly------------------ProGlnGlyLeuProGlySerProGlyAlaProGlyThr 278

Qy       1499 CCCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACC 1440
              ||||||   :::|||  |||                ||||||   |||      |||  ||
Db        279 ProGlyProGlnGlyLeuPro-----------GlySerProGlyAla-ProGlyThrPr 294

Qy       1439 CGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACC-- 1382
              ||||      |||                 |||               |||  ||||||
Db        294 oGlyProGlnGlyLeuProGlySerProGly--------------AlaProGlyThrPr 309

Qy       1381 -GGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCT 1323
                |||   |||||||||  ||||||      |||
Db        309 oGly--ProGlnGly---LeuProGlySerPro-----------------------Gly 319

Qy       1322 GCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACA--- 1266
              ||||||     |||   |||:::        ||||||                ||||||
Db        320 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 339

Qy       1265 ----------------------------GCCCCAGGCTGCCCATCC---GAACGCCTT 1239
                                           |||||||||   |||      :::   |||
Db        340 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 359

Qy       1238 CATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTACAGCC 1179
                  |||        |||      |||                        ||||||:::
Db        360 ProGlySerProGlyAlaProGlyThrProGlyPro---GlnGlyLeuProGlySerPro 378

Qy       1178 CCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGC------ 1125
                 |||     |||                               :::|||   |||
Db        379 GlyAlaProGlyThrProGlyProGlnGly-----------LeuProGlySerProGly 394

Qy       1124 ------TGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACA 1071
                       |||      |||:::                    ||||||   |||:::
Db        395 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 414

Qy       1070 GCT------GGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGG--- 1020
                       |||      |||      |||   ||||||:::          |||   |||
Db        415 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 434

Qy       1019 CCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGG 960
              ||||||       :::      |||        ||||||      |||       :::
Db        435 ProGlySerProGlyAlaProGlyThr---ProGlyProGlnGlyLeuProGlySerPro 453
```

```
                   ||:::          |||||||||            |||         |||
Db        206 roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 226

Qy       1673 AGGCCCCGCAGAGCGCGGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTC 1614
                 :::|||    :::    |||      |||    :::    |||||    :::|||       |
Db        226 lySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaP 246

Qy       1613 CATTAGGGAAGGGAGCTCCAGGCTTAGGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCT 1554
              ||        |||      ||||||  ||||||             :::    |||   |||
Db        246 roGlyThrProGlyProGlnGlyLeu---ProGly-SerProGlyAlaProGlyThrPro 264

Qy       1553 CACTGCTAGCACCTCCAGTGTCCCCTCGGTATT------TGGGCAGGAACACCTGCTTCT 1500
                                      |||   |||:::           |||    |||    :::
Db        265 Gly------------------ProGlnGlyLeuProGlySerProGlyAlaProGlyThr 278

Qy       1499 CCCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACC 1440
              ||||||    :::|||   |||                ||||||    |||      |||   ||
Db        279 ProGlyProGlnGlyLeuPro-----------GlySerProGlyAla-ProGlyThrPr 294

Qy       1439 CGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACC-- 1382
              ||||      |||                 |||                |||   ||||||
Db        294 oGlyProGlnGlyLeuProGlySerProGly--------------AlaProGlyThrPr 309

Qy       1381 -GGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCT 1323
                |||  |||||||||   ||||||
Db        309 oGly--ProGlnGly---LeuProGlySerPro-----------------------Gly 319

Qy       1322 GCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACA--- 1266
              ||||||     |||    |||:::         ||||||             ||||||
Db        320 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 339

Qy       1265 -----------------------------GCCCCAGGCTGCCCATCC---GAACGCCTT 1239
                                           |||||||||   |||      :::    |||
Db        340 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 359

Qy       1238 CATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTACAGCC 1179
                   |||          |||       |||                ||||||:::
Db        360 ProGlySerProGlyAlaProGlyThrProGlyPro---GlnGlyLeuProGlySerPro 378

Qy       1178 CCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGC------ 1125
                   |||    |||                              :::|||   |||
Db        379 GlyAlaProGlyThrProGlyProGlnGly-----------LeuProGlySerProGly 394

Qy       1124 ------TGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACA 1071
                       |||     |||:::                          ||||||   |||:::
Db        395 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 414

Qy       1070 GCT------GGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGG--- 1020
                       |||     |||     |||    ||||||:::          |||    |||
Db        415 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 434

Qy       1019 CCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGG 960
              ||||||          :::      |||      ||||||      |||          :::
Db        435 ProGlySerProGlyAlaProGlyThr---ProGlyProGlnGlyLeuProGlySerPro 453

Qy        959 TGGGGCCCAGCGCTGCCT---CCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGA 903
                    |||     |||   ||||||                                   |||
Db        454 GlyAlaProGlyThrProGlyProGln-----------------------------Gly 463

Qy        902 AGATGAGGGTGAGCAGGCCAAAGAGGCACTCCT--------------CCTGGGTGCCCA 858
                    |||     |||      |||||||||                   ||||||    |||
Db        464 LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro 483

Qy        857 GGTAGGGGGCCAGGGCACTGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGC 798
              |||      ||||||         |||                      ||||||||||||:::
Db        484 Gly---AlaProGly------ThrProGlyProGlnGlyLeuProGlySerProGlyAla 500

Qy        797 CCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGT 741
              |||                                     ||||||    |||    ||||||
Db        501 Pro-------------------------GlyThrProGlyProGlnGlyLeuProGly 511

Qy        740 ---CCCGGAAGAGGTCAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGA 684
                 ||||||                |||||||                   :::||||||

Db        512 SerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGly------ 529

Qy        683 AGTCCAGCAGCCCCACGCCCAGGATGAGCAGTGCCAGCTCCAGGGGCCTGGATCCGGGC 624
                :::|||       |||    |||                :::|||||||||      ||||||
Db        530 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla------ProGly 547

Qy        623 ACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCA 564
              |||                         |||    |||            ||||||    |||    |||
Db        548 ThrProGly--------------ProGlnGly---LeuProGlySerProGlyAlaPro 561

Qy        563 AGGACAGTG---CCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCAC 507
                 |||       |||:::          |||
Db        562 GlyThrProGlyProGlnGlyLeuProGly----------------------------- 571

Qy        506 TGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCA 447
                   |||    ||||||          |||||||||                          |||
Db        572 ------SerProGlyAlaProGlyThrProGlyProGln-----------GlyLeuPro 585

Qy        446 TGGTCATGAACTTCTCCTCTACCCCCA-----------CTTCCAGCAGCAGAGGCGGCA 399
                 |||        :::|||      |||              ||||||    :::          |||
Db        586 GlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 605

Qy        398 CATAGGTGATGCCTGCGGCCAAACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCA 339
                                   |||    ||||||       |||    |||    |||    |||
Db        606 -----------------ProGlyThrProGlyProGlnGlyLeuProGlySerProGly 619

Qy        338 AGAGCTGGGCTTTCCGGT---GCCGCAGCAGGCGGCTCACCCACAGCCTCTGGACCATAG 282
                   |||        |||                 |||||||||    |||    |||
Db        620 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr-Pr 639

Qy        281 TGGGCCAGGCGGG---TAGGGCTCAGGGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCT 225
                ||||||    |||        ||||||     |||        |||||||||                ||
Db        639 oGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLe 659

Qy        224 CGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAACTGCCTAGGAATCA 165
              |        :::|||    |||         :::|||                     |||:::
Db        659 uProGlySerProGlyAlaProGlyThrProGly--------------ProGlnGlyLe 674

Qy        164 GCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGCTTCTCAGCCCATGCTCA- 106
                ||||||:::               |||            ||||||             |||    |||
Db        674 uProGlySerProGlyAla---ProGlyThrProGlyProGlnGly-LeuProGlySerP 693

Qy        105 -------------ACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTCATCACT 60
                                ||||||       |||    |||    |||
Db        693 roGlyAlaProGlyThrProGlyProGlnGly---LeuProGlySerProGlyAlaProG 712

Qy         59 CAGATCCTGGC---CGAGGCGCGCGGCTGTCACCCGGAGCC 22
                    ||||||    :::|||           ||||||||||||
Db        712 lyThrProGlyProGlnGlyLeuProGlySerProGlyAla 725
```

RESULT 38
US-08-482-085B-103
; Sequence 103, Application US/08482085B
; Patent No. 6018030
; GENERAL INFORMATION:
;   APPLICANT: Ferrari, Franco A.
;   APPLICANT: Richardson, Charles
;   APPLICANT: Chambers, James
;   APPLICANT: Causey, Stuart
;   APPLICANT: Pollock, Thomas J.
;   APPLICANT: Cappello, Joseph
;   APPLICANT: Crissman, John W.
;   TITLE OF INVENTION: No. 6018030el Peptides Comprising Repetitive
;   TITLE OF INVENTION: Units of Amino Acids and DNA Sequences Encoding the Same
;   NUMBER OF SEQUENCES: 112
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
;   STREET: Four Embarcadero Center, Suite 3400
;   CITY: San Francisco
;   STATE: California
;   COUNTRY: US
;   ZIP: 94111
;   COMPUTER READABLE FORM:

```
Qy        398 CATAGGTGATGCCTGCGGCCAAACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCA 339
                                ||| ||||||       |||  |||  |||  |||
Db        606 -----------------ProGlyThrProGlyProGlnGlyLeuProGlySerProGly 619

Qy        338 AGAGCTGGGCTTTCCGGT---GCCGCAGCAGGCGGCTCACCCACAGCCTCTGGACCATAG 282
                  |||     |||                 |||||||||  |||  |||
Db        620 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr-Pr 639

Qy        281 TGGGCCAGGCGGG---TAGGGCTCAGGGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCT 225
               ||||||   |||     ||||||  |||      |||||||||              ||
Db        639 oGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLe 659

Qy        224 CGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAACTGCCTAGGAATCA 165
              |      :::|||  |||      :::|||                 |||:::
Db        659 uProGlySerProGlyAlaProGlyThrProGly--------------ProGlnGlyLe 674

Qy        164 GCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGCTTCTCAGCCCATGCTCA- 106
               ||||||:::          |||       ||||||         |||  |||
Db        674 uProGlySerProGlyAla---ProGlyThrProGlyProGlnGly-LeuProGlySerP 693

Qy        105 -------------ACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTCATCACT 60
                          ||||||       |||   |||   |||
Db        693 roGlyAlaProGlyThrProGlyProGlnGly---LeuProGlySerProGlyAlaProG 712

Qy         59 CAGATCCTGGC---CGAGGCGCGCGGCTGTCACCCGGAGCC 22
                 |||||| . :::|||         |||||||||||
Db        712 lyThrProGlyProGlnGlyLeuProGlySerProGlyAla 725
```

RESULT 37
US-08-707-237A-75
; Sequence 75, Application US/08707237A
; Patent No. 5830713
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Capello, Joseph
; APPLICANT: Crissman, John W.
; APPLICANT: Dorman, Mary A.
; TITLE OF INVENTION: METHODS FOR PREPARING SYNTHETIC
; TITLE OF INVENTION: REPETITIVE DNA
; NUMBER OF SEQUENCES: 108
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/707,237A
; FILING DATE: 03-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/175,155
; FILING DATE: 29-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/053,049
; FILING DATE: 22-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/609,716
; FILING DATE: 06-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/269,429
; FILING DATE: 09-NOV-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/114,618
; FILING DATE: 29-OCT-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/927,258
; FILING DATE: 04-NOV-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-55186-10/WHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 75:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 837 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-707-237A-75

Alignment Scores:

Pred. No.:	2.61e-19	Length:	837
Score:	356.50	Matches:	263
Percent Similarity:	35.16%	Conservative:	32
Best Local Similarity:	31.35%	Mismatches:	327
Query Match:	5.73%	Indels:	218
DB:	2	Gaps:	57

US-09-759-143-110 (1-3410) x US-08-707-237A-75 (1-837)

```
Qy       2347 GGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCTCAA 2288
               |||   ||||||                 ||||||    |||      |||
Db         46 GlySerProGly---------------AlaProGlyThrPro------GlyProGlnGly 58

Qy       2287 CTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCAGAGTCCCCGCATTCC 2228
               |||  |||               |||  |||  |||            |||
Db         59 LeuProGly---------------SerProGlyAlaProGlyThrProGlyProGlnGly 73

Qy       2227 AGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGC 2168
                        ||||||  |||  |||            |||   |||    :::|||
Db         74 LeuProGlySerPro---GlyAlaPro---------GlyThrProGlyPro---GlnGly 88

Qy       2167 CTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCCCCAGCCCCCAGCT 2108
               |||     |||||||||                     |||  |||     |||
Db         89 Leu---ProGlySerPro--------------------GlyAlaProGlyThrProGly 100

Qy       2107 GTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGC 2048
                |||                |||             |||         |||
Db        101 ProGln-----------GlyLeuProGlySerProGlyAlaProGlyThrProGly--P 116

Qy       2047 AACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGGA 1988
                |||   |||      ||||||   ||||||                :::|||
Db        116 roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 136

Qy       1987 CCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGC 1928
                         |||   |||   ||||||
Db        136 lySerProGlyAlaProGlyThrPro--------------------------------- 144

Qy       1927 CAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCC 1868
                                |||   ||||||                      ||||||
Db        145 --------GlyProGlnGlyLeuPro--------------------GlySerProGlyA 155

Qy       1867 TGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCATAAA 1808
                     ||||||         |||            |||          ||||||
Db        155 laProGlyThrProGlyProGlnGlyLeuProGly---------------SerPro---- 168

Qy       1807 CAGGGATGGGGCCACCTGGGACAGCAGGAAGGCAC--------------------TATC 1769
                      |||    |||||||||  |||                            |
Db        169 --------GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaP 186

Qy       1768 CAGGATGGCGAGGTCCAGGCAGATGCCCCGGC------------CCGGAACCACCCTGGC 1721
               |||||      ||||||          ||||||           |||      |||   |
Db        186 roGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyP 206

Qy       1720 CTCGGTG-------GGCTCACCCACCACCACACGTACGGAGACATCACAGGCA------G 1674
```

```
Best Local Similarity:   31.35%     Mismatches:   327
Query Match:             5.73%      Indels:       218
DB:                      1          Gaps:         57

US-09-759-143-110 (1-3410) x US-08-642-255-101 (1-837)

Qy   2347 GGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCTCAA 2288
Db     46 GlySerProGly--------------AlaProGlyThrPro------GlyProGlnGly 58

Qy   2287 CTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCAGAGTCCCCGCATTCC 2228
Db     59 LeuProGly--------------SerProGlyAlaProGlyThrProGlyProGlnGly 73

Qy   2227 AGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGC 2168
Db     74 LeuProGlySerPro---GlyAlaPro---------GlyThrProGlyPro---GlnGly 88

Qy   2167 CTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCCCCAGCCCCCAGCT 2108
Db     89 Leu---ProGlySerPro--------------------GlyAlaProGlyThrProGly 100

Qy   2107 GTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGC 2048
Db    101 ProGln-----------GlyLeuProGlySerProGlyAlaProGlyThrProGly--P 116

Qy   2047 AACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGGA 1988
Db    116 roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 136

Qy   1987 CCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGC 1928
Db    136 lySerProGlyAlaProGlyThrPro---------------------------------- 144

Qy   1927 CAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCC 1868
Db    145 --------GlyProGlnGlyLeuPro-------------------GlySerProGlyA 155

Qy   1867 TGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCATAAA 1808
Db    155 laProGlyThrProGlyProGlnGlyLeuProGly--------------SerPro---- 168

Qy   1807 CAGGGATGGGGCCACCTGGGACAGCAGGAAGGCAC--------------------TATC 1769
Db    169 --------GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaP 186

Qy   1768 CAGGATGGCGAGGTCCAGGCAGATGCCCCGGC-----------CCGGAACCACCCTGGC 1721
Db    186 roGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyP 206

Qy   1720 CTCGGTG-------GGCTCACCCACCACCACACGTACGGAGACATCACAGGCA------G 1674
Db    206 roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 226

Qy   1673 AGGCCCCGCAGAGCGCGGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTC 1614
Db    226 lySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaP 246

Qy   1613 CATTAGGGAAGGGAGCTCCAGGCTTAGGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCT 1554
Db    246 roGlyThrProGlyProGlnGlyLeu---ProGly-SerProGlyAlaProGlyThrPro 264

Qy   1553 CACTGCTAGCACCTCCAGTGTCCCCTCGGTATT------TGGGCAGGAACACCTGCTTCT 1500
Db    265 Gly-----------------ProGlnGlyLeuProGlySerProGlyAlaProGlyThr 278

Qy   1499 CCCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACC 1440
Db    279 ProGlyProGlnGlyLeuPro-----------GlySerProGlyAla-ProGlyThrPr 294

Qy   1439 CGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACC-- 1382
Db    294 oGlyProGlnGlyLeuProGlySerProGly--------------AlaProGlyThrPr 309

Qy   1381 -GGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCT 1323
Db    309 oGly--ProGlnGly---LeuProGlySerPro---------------------Gly 319

Qy   1322 GCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACA--- 1266
Db    320 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 339

Qy   1265 ----------------------------GCCCCAGGCTGCCCATCC---GAACGCCTT 1239
Db    340 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 359

Qy   1238 CATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTACAGCC 1179
Db    360 ProGlySerProGlyAlaProGlyThrProGlyPro---GlnGlyLeuProGlySerPro 378

Qy   1178 CCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGC------ 1125
Db    379 GlyAlaProGlyThrProGlyProGlnGly-----------LeuProGlySerProGly 394

Qy   1124 ------TGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACA 1071
Db    395 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 414

Qy   1070 GCT------GGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGG--- 1020
Db    415 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 434

Qy   1019 CCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGG 960
Db    435 ProGlySerProGlyAlaProGlyThr---ProGlyProGlnGlyLeuProGlySerPro 453

Qy    959 TGGGGCCCAGCGCTGCCT---CCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGA 903
Db    454 GlyAlaProGlyThrProGlyProGln---------------------------Gly 463

Qy    902 AGATGAGGGTGAGCAGGCCAAAGAGGCACTCCT--------------CCTGGGTGCCCA 858
Db    464 LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro 483

Qy    857 GGTAGGGGGCCAGGGCACTGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGC 798
Db    484 Gly---AlaProGly------ThrProGlyProGlnGlyLeuProGlySerProGlyAla 500

Qy    797 CCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGT 741
Db    501 Pro------------------------GlyThrProGlyProGlnGlyLeuProGly 511

Qy    740 ---CCCGGAAGAGGTCAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGA 684
Db    512 SerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGly------ 529

Qy    683 AGTCCAGCAGCCCCACGCCCAGGATGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGC 624
Db    530 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla------ProGly 547

Qy    623 ACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCA 564
Db    548 ThrProGly--------------ProGlnGly---LeuProGlySerProGlyAlaPro 561

Qy    563 AGGACAGTG---CCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCAC 507
Db    562 GlyThrProGlyProGlnGlyLeuProGly----------------------------- 571

Qy    506 TGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCA 447
Db    572 ------SerProGlyAlaProGlyThrProGlyProGln-----------GlyLeuPro 585

Qy    446 TGGTCATGAACTTCTCCTCTACCCCCA------------CTTCCAGCAGCAGAGGCGGCA 399
Db    586 GlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 605
```

```
                 |||  |||||||||   ||||||       |||
Db      309 oGly--ProGlnGly---LeuProGlySerPro-----------------------Gly 319

Qy     1322 GCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACA--- 1266
             ||||||      |||   |||:::      ||||||           ||||||
Db      320 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 339

Qy     1265 ------------------------------GCCCCAGGCTGCCCATCC---GAACGCCTT 1239
                                          |||||||||  |||    :::   |||
Db      340 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 359

Qy     1238 CATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTACAGCC 1179
                    |||        |||     |||                |||||||:::
Db      360 ProGlySerProGlyAlaProGlyThrProGlyPro---GlnGlyLeuProGlySerPro 378

Qy     1178 CCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGC------ 1125
                 |||        |||                          :::|||    |||
Db      379 GlyAlaProGlyThrProGlyProGlnGly-----------LeuProGlySerProGly 394

Qy     1124 ------TGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACA 1071
                     |||       |||:::                 ||||||   |||:::
Db      395 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 414

Qy     1070 GCT------GGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGG--- 1020
                     |||       |||       |||    ||||||:::      |||   |||
Db      415 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 434

Qy     1019 CCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGG 960
             ||||||         :::        |||       ||||||     |||       :::
Db      435 ProGlySerProGlyAlaProGlyThr---ProGlyProGlnGlyLeuProGlySerPro 453

Qy      959 TGGGGCCCAGCGCTGCCT---CCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGA 903
                  |||       |||   ||||||                                |||
Db      454 GlyAlaProGlyThrProGlyProGln-----------------------------Gly 463

Qy      902 AGATGAGGGTGAGCAGGCCAAAGAGGCACTCCT-----------------CCTGGGTGCCCA 858
                  |||       |||     |||||||||                    ||||||    |||
Db      464 LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro 483

Qy      857 GGTAGGGGGCCAGGGCACTGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGC 798
             |||       ||||||            |||               ||||||||||||:::
Db      484 Gly---AlaProGly------ThrProGlyProGlnGlyLeuProGlySerProGlyAla 500

Qy      797 CCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGT 741
             |||                                     ||||||   |||   ||||||
Db      501 Pro-------------------------GlyThrProGlyProGlnGlyLeuProGly 511

Qy      740 ---CCCGGAAGAGGTCAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGA 684
                ||||||                    ||||||              :::||||||
Db      512 SerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGly------ 529

Qy      683 AGTCCAGCAGCCCCACGCCCAGGATGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGC 624
             :::|||       |||   |||               :::|||||||||      ||||||
Db      530 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla------ProGly 547

Qy      623 ACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCA 564
             |||                           |||   |||         ||||||   |||   |||
Db      548 ThrProGly--------------ProGlnGly---LeuProGlySerProGlyAlaPro 561

Qy      563 AGGACAGTG---CCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCAC 507
                |||     |||:::            |||
Db      562 GlyThrProGlyProGlnGlyLeuProGly----------------------------- 571

Qy      506 TGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCA 447
                  |||   ||||||         |||||||||                          |||
Db      572 ------SerProGlyAlaProGlyThrProGlyProGln-----------GlyLeuPro 585

Qy      446 TGGTCATGAACTTCTCCTCTACCCCCA-----------CTTCCAGCAGCAGAGGCGGCA 399
             |||       :::|||      |||             ||||||    :::      |||
Db      586 GlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 605

Qy      398 CATAGGTGATGCCTGCGGCCAAACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCA 339
                          |||   ||||||         |||    |||    |||    |||
Db      606 ------------------ProGlyThrProGlyProGlnGlyLeuProGlySerProGly 619

Qy      338 AGAGCTGGGCTTTCCGGT---GCCGCAGCAGGCGGCTCACCCACAGCCTCTGGACCATAG 282
                  |||       |||                ||||||||   |||    |||
Db      620 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr-Pr 639

Qy      281 TGGGCCAGGCGGG---TAGGGCTCAGGGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCT 225
               ||||||   |||         ||||||   |||        |||||||||           ||
Db      639 oGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLe 659

Qy      224 CGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAACTGCCTAGGAATCA 165
              |      :::|||    |||          :::|||                   |||:::
Db      659 uProGlySerProGlyAlaProGlyThrProGly-------------ProGlnGlyLe 674

Qy      164 GCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGCTTCTCAGCCCATGCTCA- 106
              ||||||:::            |||           ||||||          |||    |||
Db      674 uProGlySerProGlyAla---ProGlyThrProGlyProGlnGly-LeuProGlySerP 693

Qy      105 -------------ACACCTGCTGCTGTGGGCACCTCAGTGGGGACACGTCTCATCACT 60
                         ||||||            |||    |||    |||
Db      693 roGlyAlaProGlyThrProGlyProGlnGly---LeuProGlySerProGlyAlaProG 712

Qy       59 CAGATCCTGGC---CGAGGCGCGCGGCTGTCACCCGGAGCC 22
                  ||||||   :::|||         |||||||||||
Db      712 lyThrProGlyProGlnGlyLeuProGlySerProGlyAla 725
```

RESULT 36
US-08-642-255-101
; Sequence 101, Application US/08642255
; Patent No. 5773249
; GENERAL INFORMATION:
; APPLICANT: CAPPELLO, Joseph
; APPLICANT: FERRARI, Franco A.
; TITLE OF INVENTION: High Molecular Weight Collagen-Like
; TITLE OF INVENTION: Protein Polymers
; NUMBER OF SEQUENCES: 135
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/642,255
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ROWLAND, Bertram I.
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A55556-3/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 101:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 837 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-642-255-101

Alignment Scores:

Pred. No.:	2.61e-19	Length:	837
Score:	356.50	Matches:	263
Percent Similarity:	35.16%	Conservative:	32

```
                 ||||||:::         |||          ||||||        |||  |||
Db        674 uProGlySerProGlyAla---ProGlyThrProGlyProGlnGly-LeuProGlySerP 693

Qy        105 --------------ACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTCATCACT 60
                            ||||||         |||  |||  |||
Db        693 roGlyAlaProGlyThrProGlyProGlnGly---LeuProGlySerProGlyAlaProG 712

Qy         59 CAGATCCTGGC---CGAGGCGCGCGGCTGTCACCCGGAGCC 22
                  ||||||  :::|||        |||||||||||
Db        712 lyThrProGlyProGlnGlyLeuProGlySerProGlyAla 725

RESULT 35
US-08-477-509B-103
; Sequence 103, Application US/08477509B
; Patent No. 5770697
;  GENERAL INFORMATION:
;    APPLICANT:  Ferrari, Franco A
;    APPLICANT:  Cappello, Joseph
;    APPLICANT:  Crissman, John w
;    APPLICANT:  Dorman, Mary A
;    TITLE OF INVENTION:  No. 5770697el Peptides Comprising Repetitive
;    TITLE OF INVENTION:  Units of Amino Acids and DNA Sequences Encoding the Same
;    NUMBER OF SEQUENCES:  112
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herbert
;      STREET:  Four Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/477,509B
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/175,155
;      FILING DATE:  29-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/053,049
;      FILING DATE:  22-APR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/114,618
;      FILING DATE:  29-OCT-1987
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 06/927,258
;      FILING DATE:  04-NOV-1986
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Trecartin, Richard F.
;      REGISTRATION NUMBER:  31,801
;      REFERENCE/DOCKET NUMBER:  A-55186-7/RFT/MTK
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-781-1989
;      TELEFAX:  415-398-3249
;  INFORMATION FOR SEQ ID NO:  103:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  837 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-477-509B-103

Alignment Scores:
Pred. No.:              2.61e-19       Length:          837
Score:                  356.50         Matches:         263
Percent Similarity:     35.16%         Conservative:    32
Best Local Similarity:  31.35%         Mismatches:      327
Query Match:            5.73%          Indels:          218
DB:                     1              Gaps:            57

US-09-759-143-110 (1-3410) x US-08-477-509B-103 (1-837)

Qy       2347 GGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCTCAA 2288
              |||  ||||||                  ||||||   |||     |||
Db         46 GlySerProGly---------------AlaProGlyThrPro------GlyProGlnGly 58

Qy       2287 CTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCAGAGTCCCCGCATTCC 2228
              |||  |||                |||  |||  |||           |||
Db         59 LeuProGly---------------SerProGlyAlaProGlyThrProGlyProGlnGly 73

Qy       2227 AGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGC 2168
                      ||||||  |||  |||          |||      |||   ::::|||
Db         74 LeuProGlySerPro---GlyAlaPro---------GlyThrProGlyPro---GlnGly 88

Qy       2167 CTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCCCCAGCCCCCAGCT 2108
              |||      |||||||||                        |||   |||   |||
Db         89 Leu---ProGlySerPro--------------------GlyAlaProGlyThrProGly 100

Qy       2107 GTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGC 2048
                |||                 |||                |||          |||
Db        101 ProGln------------GlyLeuProGlySerProGlyAlaProGlyThrProGly--P 116

Qy       2047 AACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGGA 1988
                |||   |||   ||||||  ||||||                 :::|||
Db        116 roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 136

Qy       1987 CCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGC 1928
                     |||  |||  ||||||
Db        136 lySerProGlyAlaProGlyThrPro-------------------------------- 144

Qy       1927 CAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCC 1868
                              |||  ||||||                       ||||||
Db        145 --------GlyProGlnGlyLeuPro-------------------GlySerProGlyA 155

Qy       1867 TGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCATAAA 1808
                    ||||||   |||      |||                 ||||||
Db        155 laProGlyThrProGlyProGlnGlyLeuProGly---------------SerPro---- 168

Qy       1807 CAGGGATGGGGCCACCTGGGACAGCAGGAAGGCAC--------------------TATC 1769
                      |||  |||||||||   |||                           |
Db        169 --------GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaP 186

Qy       1768 CAGGATGGCGAGGTCCAGGCAGATGCCCCGGC------------CCGGAACCACCCTGGC 1721
              |||||      ||||||          ||||||       |||      |||   |
Db        186 roGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyP 206

Qy       1720 CTCGGTG-------GGCTCACCCACCACCACACGTACGGAGACATCACAGGCA------G 1674
              ||:::            |||||||||      |||         |||
Db        206 roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 226

Qy       1673 AGGCCCCGCAGAGCGCGGGTGGAGGTGGAGCAGGCCACTGCCTCCAGCACCCACGTGTC 1614
                :::|||   :::   |||       |||    :::   ||||||   :::|||       |
Db        226 lySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaP 246

Qy       1613 CATTAGGGAAGGGAGCTCCAGGCTTAGGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCT 1554
              ||       |||       ||||||   ||||||           :::   |||   |||
Db        246 roGlyThrProGlyProGlnGlyLeu---ProGly-SerProGlyAlaProGlyThrPro 264

Qy       1553 CACTGCTAGCACCTCCAGTGTCCCCTCGGTATT------TGGGCAGGAACACCTGCTTCT 1500
                                |||  |||:::               |||   |||   :::
Db        265 Gly-----------------ProGlnGlyLeuProGlySerProGlyAlaProGlyThr 278

Qy       1499 CCCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACC 1440
              ||||||   :::|||  |||                ||||||  |||      |||   ||
Db        279 ProGlyProGlnGlyLeuPro-----------GlySerProGlyAla-ProGlyThrPr 294

Qy       1439 CGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACC-- 1382
              ||||      |||                 |||                |||  ||||||
Db        294 oGlyProGlnGlyLeuProGlySerProGly---------------AlaProGlyThrPr 309

Qy       1381 -GGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCT 1323
```

```
Qy   2167 CTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCCCCAGCCCCCAGCT 2108
Db     89 Leu---ProGlySerPro--------------------GlyAlaProGlyThrProGly 100

Qy   2107 GTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGC 2048
Db    101 ProGln-----------GlyLeuProGlySerProGlyAlaProGlyThrProGly--P 116

Qy   2047 AACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGGA 1988
Db    116 roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 136

Qy   1987 CCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGC 1928
Db    136 lySerProGlyAlaProGlyThrPro-------------------------------- 144

Qy   1927 CAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCC 1868
Db    145 --------GlyProGlnGlyLeuPro-------------------GlySerProGlyA 155

Qy   1867 TGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCATAAA 1808
Db    155 laProGlyThrProGlyProGlnGlyLeuProGly---------------SerPro---- 168

Qy   1807 CAGGGATGGGGCCACCTGGGACAGCAGGAAGGCAC--------------------TATC 1769
Db    169 --------GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaP 186

Qy   1768 CAGGATGGCGAGGTCCAGGCAGATGCCCCGGC------------CCGGAACCACCCTGGC 1721
Db    186 roGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyP 206

Qy   1720 CTCGGTG-------GGCTCACCCACCACCACACGTACGGAGACATCACAGGCA------G 1674
Db    206 roGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProG 226

Qy   1673 AGGCCCCGCAGAGCGCGGGTGGAGGTGGAGCAGGCCACTGCCTCCAGCACCCACGTGTC 1614
Db    226 lySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaP 246

Qy   1613 CATTAGGGAAGGGAGCTCCAGGCTTAGGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCT 1554
Db    246 roGlyThrProGlyProGlnGlyLeu---ProGly-SerProGlyAlaProGlyThrPro 264

Qy   1553 CACTGCTAGCACCTCCAGTGTCCCCTCGGTATT------TGGGCAGGAACACCTGCTTCT 1500
Db    265 Gly-----------------ProGlnGlyLeuProGlySerProGlyAlaProGlyThr 278

Qy   1499 CCCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACC 1440
Db    279 ProGlyProGlnGlyLeuPro------------GlySerProGlyAla-ProGlyThrPr 294

Qy   1439 CGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACC-- 1382
Db    294 oGlyProGlnGlyLeuProGlySerProGly---------------AlaProGlyThrPr 309

Qy   1381 -GGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCT 1323
Db    309 oGly--ProGlnGly---LeuProGlySerPro-----------------------Gly 319

Qy   1322 GCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACA--- 1266
Db    320 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 339

Qy   1265 ------------------------------GCCCCAGGCTGCCCATCC---GAACGCCTT 1239
Db    340 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 359

Qy   1238 CATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTACAGCC 1179
Db    360 ProGlySerProGlyAlaProGlyThrProGlyPro---GlnGlyLeuProGlySerPro 378

Qy   1178 CCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGC------ 1125
Db    379 GlyAlaProGlyThrProGlyProGlnGly------------LeuProGlySerProGly 394

Qy   1124 ------TGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACA 1071
Db    395 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 414

Qy   1070 GCT------GGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGG--- 1020
Db    415 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 434

Qy   1019 CCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGG 960
Db    435 ProGlySerProGlyAlaProGlyThr---ProGlyProGlnGlyLeuProGlySerPro 453

Qy    959 TGGGGCCCAGCGCTGCCT---CCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGA 903
Db    454 GlyAlaProGlyThrProGlyProGln----------------------------Gly 463

Qy    902 AGATGAGGGTGAGCAGGCCAAAGAGGCACTCCT---------------CCTGGGTGCCCA 858
Db    464 LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro 483

Qy    857 GGTAGGGGGCCAGGGCACTGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGC 798
Db    484 Gly---AlaProGly------ThrProGlyProGlnGlyLeuProGlySerProGlyAla 500

Qy    797 CCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGT 741
Db    501 Pro-------------------------GlyThrProGlyProGlnGlyLeuProGly 511

Qy    740 ---CCCGGAAGAGGTCAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGA 684
Db    512 SerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGly------ 529

Qy    683 AGTCCAGCAGCCCCACGCCCAGGATGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGC 624
Db    530 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla------ProGly 547

Qy    623 ACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCA 564
Db    548 ThrProGly---------------ProGlnGly---LeuProGlySerProGlyAlaPro 561

Qy    563 AGGACAGTG---CCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCAC 507
Db    562 GlyThrProGlyProGlnGlyLeuProGly---------------------------- 571

Qy    506 TGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCA 447
Db    572 ------SerProGlyAlaProGlyThrProGlyProGln------------GlyLeuPro 585

Qy    446 TGGTCATGAACTTCTCCTCTACCCCCA------------CTTCCAGCAGCAGAGGCGGCA 399
Db    586 GlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 605

Qy    398 CATAGGTGATGCCTGCGGCCAAACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCA 339
Db    606 ------------------ProGlyThrProGlyProGlnGlyLeuProGlySerProGly 619

Qy    338 AGAGCTGGGCTTTCCGGT---GCCGCAGCAGGCGGCTCACCCACAGCCTCTGGACCATAG 282
Db    620 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr-Pr 639

Qy    281 TGGGCCAGGCGGG---TAGGGCTCAGGGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCT 225
Db    639 oGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLe 659

Qy    224 CGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAACTGCCTAGGAATCA 165
Db    659 uProGlySerProGlyAlaProGlyThrProGly---------------ProGlnGlyLe 674

Qy    164 GCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGCTTCTCAGCCCATGCTCA- 106
```

```
Qy     1106 AGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACAGCT------GGTGCAGCCGGGGAA 1053
Db      381 GlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySer 400

Qy     1052 GCAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGG---CCCGGCATGGACAGCAGTGGGGCG 996
Db      401 ProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGly 420

Qy      995 ACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCT---CCT 939
Db      421 Thr---ProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro 439

Qy      938 CAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGA 879
Db      440 Gln----------------------------GlyLeuProGlySerProGlyAlaPro 449

Qy      878 GGCACTCCT---------------CCTGGGTGCCCAGGTAGGGGGCCAGGGCACTGGTGT 834
Db      450 GlyThrProGlyProGlnGlyLeuProGlySerProGly---AlaProGly------Thr 466

Qy      833 CCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCAT 774
Db      467 ProGlyProGlnGlyLeuProGlySerProGlyAlaPro-------------------- 479

Qy      773 AGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGT---CCCGGAAGAGGTCAGAGAGCA 720
Db      480 ------GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 497

Qy      719 GGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGCCCAGGA 660
Db      498 GlyProGlnGlyLeuProGlySerProGly------AlaProGlyThrProGlyProGln 515

Qy      659 TGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGG 600
Db      516 GlyLeuProGlySerProGlyAla------ProGlyThrProGly-------------- 528

Qy      599 CCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCAAGGACAGTG---CCCAGATGAAGG 543
Db      529 ProGlnGly---LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 547

Qy      542 GCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTAGGAGCGGGACAC 483
Db      548 ProGly-----------------------------------SerProGlyAlaProGly 555

Qy      482 AGACCAGGCCCAGCACTGGACCAATGCCCAGCACCATGGTCATGAACTTCTCCTCTACCC 423
Db      556 ThrProGlyProGln-----------GlyLeuProGlySerProGlyAlaProGlyThr 571

Qy      422 CCA-----------CTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCCAAAC 375
Db      572 ProGlyProGlnGlyLeuProGlySerProGlyAla-----------------ProGly 585

Qy      374 ACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGT---GCC 318
Db      586 ThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGln 605

Qy      317 GCAGCAGGCGGCTCACCCACAGCCTCTGGACCATAGTGGGCCAGGCGGG---TAGGGCTC 261
Db      606 GlyLeuProGlySerProGlyAlaProGlyThr-ProGlyProGlnGlyLeuProGlySe 625

Qy      260 AGGGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGC 201
Db      625 rProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGl 645

Qy      200 CTCTCCTCCTTGCTGCCGCCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCC 141
Db      645 yThrProGly--------------ProGlnGlyLeuProGlySerProGlyAla---Pr 659

Qy      140 CTTTGGTGCCGGTCCAGCTTCTCAGCCCATGCTCA--------------ACACCTGCTG 96
Db      659 oGlyThrProGlyProGlnGly-LeuProGlySerProGlyAlaProGlyThrProGlyP 679

Qy       95 CTGTGGGGCACCTCAGTGGGGACACGTCTCATCACTCAGATCCTGGC---CGAGGCGCGC 39
Db      679 roGlnGly---LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuP 698

Qy       38 GGCTGTCACCCGGAGCC 22
Db      698 roGlySerProGlyAla 703
```

```
RESULT 34
US-08-175-155-68
; Sequence 68, Application US/08175155
; Patent No. 5641648
;  GENERAL INFORMATION:
;    APPLICANT:  Ferrari, Franco A.
;    APPLICANT:  Cappello, Joseph
;    APPLICANT:  Crissman, John W.
;    APPLICANT:  Dorman, Mary A.
;    TITLE OF INVENTION:  Methods for Preparing Synthetic
;    TITLE OF INVENTION:  Repetitive DNA
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herbert
;      STREET:  Four Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/175,155
;      FILING DATE:  29-DEC-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rowland, Bertram I.
;      REGISTRATION NUMBER:  20015
;      REFERENCE/DOCKET NUMBER:  A-55186-5/BIR
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-781-1989
;      TELEFAX:  415-398-3249
;  INFORMATION FOR SEQ ID NO:  68:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  837 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-175-155-68

Alignment Scores:
Pred. No.:               2.61e-19      Length:          837
Score:                   356.50        Matches:         263
Percent Similarity:      35.16%        Conservative:    32
Best Local Similarity:   31.35%        Mismatches:      327
Query Match:             5.73%         Indels:          218
DB:                      1             Gaps:            57

US-09-759-143-110 (1-3410) x US-08-175-155-68 (1-837)

Qy     2347 GGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCTCAA 2288
Db       46 GlySerProGly--------------AlaProGlyThrPro------GlyProGlnGly 58

Qy     2287 CTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCAGAGTCCCCGCATTCC 2228
Db       59 LeuProGly--------------SerProGlyAlaProGlyThrProGlyProGlnGly 73

Qy     2227 AGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGC 2168
Db       74 LeuProGlySerPro---GlyAlaPro---------GlyThrProGlyPro---GlnGly 88
```

```
                  |||    |||    |||                          ||||||    :::|||
Db        679 roGlnGly---LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuP 698

Qy         38 GGCTGTCACCCGGAGCC 22
                  ||||||||||||
Db        698 roGlySerProGlyAla 703

RESULT 33
US-08-397-633A-53
; Sequence 53, Application US/08397633A
; Patent No. 5773577
;  GENERAL INFORMATION:
;    APPLICANT:  Cappello, Joseph
;    TITLE OF INVENTION:  PRODUCTS COMPRISING SUBSTRATESCAPABLE
;    TITLE OF INVENTION:  OF ENZYMATIC CROSS-LINKING
;    NUMBER OF SEQUENCES:  105
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/397,633A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rowland, Bertram I
;      REGISTRATION NUMBER:  20,015
;      REFERENCE/DOCKET NUMBER:  A-58848-1/BIR PROP-011-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 781-1989
;      TELEFAX:  (415) 398-3249
;      TELEX:  910 277299
;  INFORMATION FOR SEQ ID NO:  53:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  829 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-397-633A-53

Alignment Scores:
Pred. No.:               2.38e-19     Length:          829
Score:                   357.00       Matches:         256
Percent Similarity:      34.78%       Conservative:    33
Best Local Similarity:   30.81%       Mismatches:      324
Query Match:             5.73%        Indels:          219
DB:                      1            Gaps:            53

US-09-759-143-110 (1-3410) x US-08-397-633A-53 (1-829)

Qy       2320 GCTCCCAAAAACCCTTCTCTAGGTGTGTCT---CAACTAGGAGGCTAGCTGTTAACCCTG 2264
              ||||||       |||      |||              |||||||||
Db         32 AlaProGlyThrProGlyGluGlyGlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGly 51

Qy       2263 AGCCTGGGTAATCCACCTGCAGAGTCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTC 2204
                 :::|||:::|||   |||                          |||   ||||||
Db         52 AspValGlySerProGlyAlaProGlyThrPro------------GlyProGlnGlyLeu 67

Qy       2203 CCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGAC 2144
              |||                                         |||||||||
Db         68 Pro------------------------------------------GlySerPro------ 71

Qy       2143 TGGGGAGAGAGGAGAGGGACGCCCCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCA 2084
                               |||    |||       |||       |||
Db         72 ---------------GlyAlaProGlyThrProGlyProGln-----------GlyLeu 82

Qy       2083 CAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGG 2024
                 |||      |||      |||      |||   |||   |||      ||||||
Db         83 ProGlySerProGlyAlaProGlyThrProGly--ProGlnGlyLeuProGlySerProG 102

Qy       2023 CAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGGACCCAGTGAGGCAGGCCCTCCACCC 1964
                ||||||                  :::|||                |||   |||   ||||
Db        102 lyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrP 122

Qy       1963 CAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTAC 1904
              ||                                                |||   ||||
Db        122 ro------------------------------------------GlyProGlnGlyLeuP 128

Qy       1903 CTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGT 1844
              ||                                   ||||||         ||||||   |||
Db        128 ro--------------------GlySerProGlyAlaProGlyThrProGlyProGlnG 141

Qy       1843 GACAGACTGGCTGAGCTGGACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAG 1784
                    |||               ||||||              |||   |||||||||
Db        141 lyLeuProGly---------------SerPro------------GlyAlaProGlyThrP 152

Qy       1783 CAGGAAGGCAC--------------------TATCCAGGATGGCGAGGTCCAGGCAGAT 1745
                |||                                  ||||||      ||||||
Db        152 roGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyL 172

Qy       1744 GCCCCGGC-----------CCGGAACCACCCTGGCCTCGGTG-------GGCTCACCCA 1704
               ||||||              |||      |||      |||   |||:::              |||||||||
Db        172 euProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProG 192

Qy       1703 CCACCACACGTACGGAGACATCACAGGCA------GAGGCCCCGCAGAGCGCGGGTGGAG 1650
                      |||      |||                  :::|||   :::   |||
Db        192 lyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrP 212

Qy       1649 GTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGGGAAGGGAGCTCCAGGCT 1590
                |||   :::   ||||||   :::|||         |||            |||      ||||
Db        212 roGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyL 232

Qy       1589 TAGGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCTAGCACCTCCAGTGTCCC 1530
              ||   ||||||              :::    |||    |||                  |||
Db        232 eu---ProGly-SerProGlyAlaProGlyThrProGly-----------------Pro 244

Qy       1529 CTCGGTATT------TGGCAGGAACACCTGCTTCTCCCGGTGGTAGAGGGAGGCCAGTG 1476
              |||:::             |||   |||      :::||||||   :::|||   |||
Db        245 GlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuPro--- 263

Qy       1475 TGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAGCTGTCA 1416
                       ||||||   |||         |||   ||||||         |||
Db        264 ---------GlySerProGlyAla-ProGlyThrProGlyProGlnGlyLeuProGlySe 280

Qy       1415 CCACGGCCACACTGTGGGACAGGCATGTGGCACC---GGCAGCCACAGGGAAAGCTGCCA 1359
                |||                  |||   ||||||   |||   |||||||||   ||||||
Db        280 rProGly---------------AlaProGlyThrProGly--ProGlnGly---LeuPro 293

Qy       1358 CACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCTGCACCAGCCGGTCCATGACCAGAG 1299
                |||                            ||||||      |||   |||:::
Db        294 GlySerPro-----------------------GlyAlaProGlyThrProGlyProGln 305

Qy       1298 AGAAGACCAGGGAGATGGCGCACTGCAGGAACA------------------------- 1266
                 ||||||                ||||||
Db        306 GlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySer 325

Qy       1265 ------GCCCCAGGCTGCCCATCC---GAACGCCTTCATCATAGTGTCTCCGGGCCTCGG 1215
                 |||||||||   |||      :::   |||      |||            |||
Db        326 ProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGly 345

Qy       1214 TGCCCGGCTCAGCTCTGGGCACGCCCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAA 1155
                |||                  ||||||:::               |||      |||
Db        346 ThrProGlyPro---GlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro 364

Qy       1154 ACAGCGTGAAGGTCATGAGTGCCATCCAGC------------TGCACAGCTCAGCCACGA 1107
                                   :::|||   |||                  |||      |||:::
Db        365 GlnGly------------LeuProGlySerProGlyAlaProGlyThrProGlyProGln 380
```

```
Qy   2083 CAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGG 2024
              |||          |||          |||    |||   |||       ||||||
Db     83 ProGlySerProGlyAlaProGlyThrProGly--ProGlnGlyLeuProGlySerProG 102

Qy   2023 CAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGACCCAGTGAGGCAGGCCCTCCACCC 1964
            ||||||                  :::|||                |||   |||   ||||
Db    102 lyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrP 122

Qy   1963 CAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTAC 1904
           ||                                                  |||   ||||
Db    122 ro---------------------------------------GlyProGlnGlyLeuP 128

Qy   1903 CTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGT 1844
           ||                          ||||||         ||||||         |||
Db    128 ro---------------------GlySerProGlyAlaProGlyThrProGlyProGlnG 141

Qy   1843 GACAGACTGGCTGAGCTGGACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAG 1784
                  |||              ||||||           |||    |||||||||
Db    141 lyLeuProGly--------------SerPro-----------GlyAlaProGlyThrP 152

Qy   1783 CAGGAAGGCAC--------------------TATCCAGGATGGCGAGGTCCAGGCAGAT 1745
            |||                               ||||||      ||||||
Db    152 roGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyL 172

Qy   1744 GCCCCGGC-----------CCGGAACCACCCTGGCCTCGGTG-------GGCTCACCCA 1704
            ||||||          |||     |||    |||:::         ||||||||
Db    172 euProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProG 192

Qy   1703 CCACCACACGTACGGAGACATCACAGGCA------GAGGCCCCGCAGAGCGCGGGTGGAG 1650
                   |||         |||           :::|||   :::   |||
Db    192 lyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrP 212

Qy   1649 GTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGGGAAGGGAGCTCCAGGCT 1590
            |||   :::   ||||||   :::|||        |||          |||        ||||
Db    212 roGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyL 232

Qy   1589 TAGGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCTAGCACCTCCAGTGTCCC 1530
           ||   ||||||          :::   |||   |||                  |||
Db    232 eu---ProGly-SerProGlyAlaProGlyThrProGly-----------------Pro 244

Qy   1529 CTCGGTATT------TGGGCAGGAACACCTGCTTCTCCCGGTGGTAGAGGGAGGCCAGTG 1476
             |||:::             |||    |||    :::||||||    :::|||   |||
Db    245 GlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuPro--- 263

Qy   1475 TGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAGCTGTCA 1416
                    ||||||     |||   |||   ||||||   |||
Db    264 ---------GlySerProGlyAla-ProGlyThrProGlyProGlnGlyLeuProGlySe 280

Qy   1415 CCACGGCCACACTGTGGGACAGGCATGTGGCACC---GGCAGCCACAGGGAAAGCTGCCA 1359
             |||               |||   ||||||   |||   |||||||||   ||||||
Db    280 rProGly--------------AlaProGlyThrProGly--ProGlnGly---LeuPro 293

Qy   1358 CACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCTGCACCAGCCGGTCCATGACCAGAG 1299
                |||                            ||||||   |||   |||:::
Db    294 GlySerPro----------------------GlyAlaProGlyThrProGlyProGln 305

Qy   1298 AGAAGACCAGGGAGATGGCGCACTGCAGGAACA-------------------------- 1266
                ||||||                  ||||||
Db    306 GlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySer 325

Qy   1265 ------GCCCCAGGCTGCCCATCC---GAACGCCTTCATCATAGTGTCTCCGGGCCTCGG 1215
                ||||||||   |||          :::   |||      |||          |||
Db    326 ProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGly 345

Qy   1214 TGCCCGGCTCAGCTCTGGGCACGCCCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAA 1155
             |||                   ||||||:::            |||      |||
Db    346 ThrProGlyPro---GlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro 364

Qy   1154 ACAGCGTGAAGGTCATGAGTGCCATCCAGC-----------TGCACAGCTCAGCCACGA 1107
                         :::|||   |||                |||     |||:::
Db    365 GlnGly-----------LeuProGlySerProGlyAlaProGlyThrProGlyProGln 380

Qy   1106 AGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACAGCT------GGTGCAGCCGGGGAA 1053
                            ||||||   |||:::                |||    |||
Db    381 GlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySer 400

Qy   1052 GCAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGG---CCCGGCATGGACAGCAGTGGGGCG 996
              |||   ||||||:::      |||   |||   ||||||            :::
Db    401 ProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGly 420

Qy    995 ACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCT---CCT 939
           |||         ||||||      |||        :::            |||    |||   |||
Db    421 Thr---ProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro 439

Qy    938 CAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGA 879
           |||                                |||        |||        |||
Db    440 Gln----------------------------GlyLeuProGlySerProGlyAlaPro 449

Qy    878 GGCACTCCT---------------CCTGGGTGCCCAGGTAGGGGGCCAGGGCACTGGTGT 834
           |||||||||                 ||||||    ||||||       ||||||
Db    450 GlyThrProGlyProGlnGlyLeuProGlySerProGly---AlaProGly------Thr 466

Qy    833 CCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCAT 774
           |||                ||||||||||||:::|||
Db    467 ProGlyProGlnGlyLeuProGlySerProGlyAlaPro-------------------- 479

Qy    773 AGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGT---CCCGGAAGAGGTCAGAGAGCA 720
                       ||||||       |||    ||||||    ||||||
Db    480 ------GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPro 497

Qy    719 GGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGCCCAGGA 660
           ||||||                :::||||||          :::|||          |||   |||
Db    498 GlyProGlnGlyLeuProGlySerProGly------AlaProGlyThrProGlyProGln 515

Qy    659 TGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGG 600
                        :::|||||||||          |||||||||
Db    516 GlyLeuProGlySerProGlyAla------ProGlyThrProGly-------------- 528

Qy    599 CCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCAAGGACAGTG---CCCAGATGAAGG 543
           |||   |||           ||||||   |||   |||    |||   |||:::
Db    529 ProGlnGly---LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu 547

Qy    542 GCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTAGGAGCGGGACAC 483
              |||                                             |||   ||||||
Db    548 ProGly------------------------------SerProGlyAlaProGly 555

Qy    482 AGACCAGGCCCAGCACTGGACCAATGCCCAGCACCATGGTCATGAACTTCTCCTCTACCC 423
              |||||||||                           |||    |||           :::|||
Db    556 ThrProGlyProGln------------GlyLeuProGlySerProGlyAlaProGlyThr 571

Qy    422 CCA------------CTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCCAAAC 375
           |||                  ||||||    :::    |||                        |||
Db    572 ProGlyProGlnGlyLeuProGlySerProGlyAla-------------------ProGly 585

Qy    374 ACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGT---GCC 318
           ||||||      |||    |||    |||    |||             |||        |||
Db    586 ThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGln 605

Qy    317 GCAGCAGGCGGCTCACCCACAGCCTCTGGACCATAGTGGGCCAGGCGGG---TAGGGCTC 261
                      |||||||||    |||    |||          ||||||    |||          |||||
Db    606 GlyLeuProGlySerProGlyAlaProGlyThr-ProGlyProGlnGlyLeuProGlySe 625

Qy    260 AGGGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGC 201
           |    |||      |||||||||                 |||          :::|||    |||
Db    625 rProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGl 645

Qy    200 CTCTCCTCCTTGCTGCCGCCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCC 141
             :::|||                         |||:::    ||||||:::             ||
Db    645 yThrProGly---------------ProGlnGlyLeuProGlySerProGlyAla---Pr 659

Qy    140 CTTTGGTGCCGGTCCAGCTTCTCAGCCCATGCTCA---------------ACACCTGCTG 96
           |           ||||||          |||    |||                    ||||||
Db    659 oGlyThrProGlyProGlnGly-LeuProGlySerProGlyAlaProGlyThrProGlyP 679

Qy     95 CTGTGGGGCACCTCAGTGGGGACACGTCTCATCACTCAGATCCTGGC---CGAGGCGCGC 39
```

```
Db        390 oGlyProProGlyAlaProGlyProAlaGlyProProGlySerArgGlyAspProGlyPr 410

Qy        981 GACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCT------------------CC 940
              |                        |||||||||  |||                  ||
Db        410 oProGlyAlaPro-----------GlyProAlaGlyProProGlySerArgGlyAspPr 426

Qy        939 TCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAG 880
              |  |||||                               |||     |||   ||||||
Db        426 oGlyProPro-----------------------GlyAlaProGlyProAlaGlyProPr 438

Qy        879 AGGCACT--------------CCTCCTGGGTGCCCAGGTAGGGGG------CCAGGGCA 841
               |||:::              |||||||||  ||||||              ||||||
Db        438 oGlySerArgGlyAspProGlyProProGlyAlaProGlyProAlaGlyProProGly-- 457

Qy        840 CTGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGACTGATCATG 781
                                    :::  |||  ||||||  |||
Db        458 ---------------------SerArgGlyAspProGlyProPro-------------- 465

Qy        780 AAGGCATAGACAGAGTAGGCCTGGCGACAGTGGTCCGGGTCCCGGAAGAGGTCAG----- 726
                                  ||||||  :::||||||         ||||||
Db        466 ------------GlyAlaProGlyProAlaGlyProProGlySerArgGlyAspProGl 481

Qy        725 ------------------AGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAG 685
                                 ||||||||||||             ||||||
Db        481 yProProGlyAlaProGlyProAlaGlyProProGlySerArgGlyAspProGly----- 499

Qy        684 AAGTCCAGCAGCCCCACGCCCAGGATGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGG 625
                 |||  ||||||  |||           |||  |||           ||||||||
Db        500 -ProProGlyAlaProGlyPro------AlaGlyProProGlySerArgGlyAspProGl 517

Qy        624 CACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAA------AGAGGCTCAGCAGG 571
              |               |||     |||   |||    :::          |||  |||||
Db        517 yProProGly------AlaHisGlyProAlaGlyProLysGlyAlaHisGlyProAlaGl 535

Qy        570 ATGCCCAAGGACAGTGCCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGG 511
              |                |||:::          |||                |||   ||
Db        535 y--------------ProLysGlyAlaHisGlyProAlaGlyProLysGlyAlaHisGl 550

Qy        510 TCACTGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGC 451
              |         |||               |||||||||||
Db        550 y---------ProAlaGlyProLysGlyAlaProGlyProAla-----------GlyPr 563

Qy        450 ACCATGGTCATGAACTTCTCCTCTACCCCCA---CTTCCAGCAGCAGAGGCGGCACATAG 394
               |||   |||           |||   ||||||      |||
Db        563 oProGlySerArgGlyAspProGlyProProGlyAlaProGly---------------- 577

Qy        393 GTGATGCCTGCGGCCAAACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCAAGAGC 334
                                |||   |||||||||  |||        |||   |||
Db        578 ------------ProAlaGlyProProGlySerArg------GlyAspProGlyProPr 591

Qy        333 TGGGCTTTCCGGTGCCGCAGCA--------------------GGCGGCTCACCCACAGC 295
               |||     |||   |||                        |||   |||   ||
Db        591 oGlyAlaProGlyProAlaGlyProProGlySerArgGlyAspProGlyProProGlyAl 611

Qy        294 CTCTGGACCATAGTGGGCCAGGCGGGTAGGGCTCAGGGGGCCGTTCAGGCACTCCAGAAC 235
              |   ||||||  :::||||||            ||||||  |||       |||   |||
Db        611 aProGlyPro-AlaGlyProPro------GlySerArgGlyAspProGlyProPro---- 627

Qy        234 TGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAACTGC 175
                                  ||||||    |||      |||
Db        628 --------------GlyAlaProGlyProAlaGlyProProGly--------------S 638

Qy        174 CTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGCTTCTCAGC 115
                |||     ||||||                |||        |||||||||
Db        638 erArgGlyAspProGly-----------ProProGlyAlaProGlyProAlaGly-Pro 653

Qy        114 CCATGCTCA--------------ACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACG 70
               |||  |||                   |||   |||   |||       :::|||
Db        654 ProGlySerArgGlyAspProGlyProProGlyAlaProGlyPro---AlaGlyProPro 672

Qy         69 TCTCATCACTCAGATCCTGGC----------------------------CGAGGCGCG 40
                          |||||||||                              |||:::
Db        673 GlySerArgGlyAspProGlyProProGlyAlaProGlyProAlaGlyProProGlySer 692

Qy         39 CGGCTGTCACCCGGA 25
                |||     ||||||
Db        693 ArgGlyAspProGly 697
```

RESULT 32
US-08-642-255-132
; Sequence 132, Application US/08642255
; Patent No. 5773249
; GENERAL INFORMATION:
; APPLICANT: CAPPELLO, Joseph
; APPLICANT: FERRARI, Franco A.
; TITLE OF INVENTION: High Molecular Weight Collagen-Like
; TITLE OF INVENTION: Protein Polymers
; NUMBER OF SEQUENCES: 135
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/642,255
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ROWLAND, Bertram I.
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A55556-3/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 132:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 829 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-642-255-132

Alignment Scores:

Pred. No.:	2.38e-19	Length:	829
Score:	357.00	Matches:	256
Percent Similarity:	34.78%	Conservative:	33
Best Local Similarity:	30.81%	Mismatches:	324
Query Match:	5.73%	Indels:	219
DB:	1	Gaps:	53

US-09-759-143-110 (1-3410) x US-08-642-255-132 (1-829)

```
Qy       2320 GCTCCCAAAAACCCTTCTCTAGGTGTGTCT---CAACTAGGAGGCTAGCTGTTAACCCTG 2264
                ||||||     |||   |||           |||||||||
Db         32 AlaProGlyThrProGlyGluGlyGlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGly 51

Qy       2263 AGCCTGGGTAATCCACCTGCAGAGTCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTC 2204
                :::|||:::|||  |||                                 |||  ||||||
Db         52 AspValGlySerProGlyAlaProGlyThrPro-----------GlyProGlnGlyLeu 67

Qy       2203 CCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGAC 2144
                |||                                        |||||||||
Db         68 Pro---------------------------------------GlySerPro------ 71

Qy       2143 TGGGGAGAGAGGAGAGGGACGCCCCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCA 2084
                               |||  |||        |||       |||
Db         72 --------------GlyAlaProGlyThrProGlyProGln-----------GlyLeu 82
```

```
;    APPLICANT:  CAPPELLO, Joseph
;    APPLICANT:  FERRARI, Franco A.
;    TITLE OF INVENTION:  High Molecular Weight Collagen-Like
;    TITLE OF INVENTION:  Protein Polymers
;    NUMBER OF SEQUENCES:  135
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/642,255
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROWLAND, Bertram I.
;      REGISTRATION NUMBER:  20,015
;      REFERENCE/DOCKET NUMBER:  A55556-3/BIR
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 494-8700
;      TELEFAX:  (415) 494-8771
;      TELEX:  910 277299 FHT UR
;  INFORMATION FOR SEQ ID NO:  62:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1064 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-642-255-62

Alignment Scores:
Pred. No.:              1.53e-19        Length:         1064
Score:                  360.00          Matches:        265
Percent Similarity:     33.26%          Conservative:   24
Best Local Similarity:  30.49%          Mismatches:     296
Query Match:            5.78%           Indels:         286
DB:                     1               Gaps:           60

US-09-759-143-110 (1-3410) x US-08-642-255-62 (1-1064)

Qy       2314 AAAAACCCTTCTCTAGGTGTGTCTCAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGT 2255
              :::||||||      |||||:::|||                 |||
Db         12 GluAsnPro------GlyValThrGln---------------LeuAsnArgLeuAlaAla 24

Qy       2254 AATCCACCTGCAGAGTCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAA 2195
              :::||||||      |||            |||       |||:::|||
Db         25 HisProProPheAlaSerAspProMetGlyAlaHisGlyProAlaGly------------ 40

Qy       2194 GTCCAGACTGAAACCCCCTTGGAAGGCCTCCAG---------------TCAGGCAGCCCT 2150
                                    :::|||                  |||:::|||
Db         41 ------------------ProLysGlyAlaHisGlyProAlaGlyProLysGlyAlaPro 54

Qy       2149 AGAGACTGGGGAGAG------AGGAGAGGGACGCCCCAGCCCCCAGCTGTGCAGCTACGC 2096
                       |||         ||||||   |||    ||||||
Db         55 GlyProAlaGlyProProGlySerArgGlyAspProGlyProPro--------------- 69

Qy       2095 ACCTCAGCAGCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGCAACAGAAACTGG 2036
                    |||   |||    :::      |||         ||||||         :::
Db         70 ------GlyAlaProGlyProAlaGlyProPro---------GlySerArgGlyAspPro 84

Qy       2035 CGGCCAGCCCGGCAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGGA------------ 1988
                 |||      |||      |||        |||      |||
Db         85 GlyProProGlyAlaProGlyProAlaGlyProProGlySerArgGlyAspProGlyPro 104

Qy       1987 ---------CCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTTCTACGCTGA 1937
                      |||        |||||||||||                    ||||||
Db        105 ProGlyAlaProGlyProAlaGlyProPro------------GlySer------------ 116

Qy       1936 GTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAG 1877
                                                                ||||||
Db        117 ---------------------------------------------*ArgGlyAspPro 120

Qy       1876 AC------------CCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGA 1830
                            ||||||          |||       :::          |||
Db        121 GlyProProGlyAlaProGlyProAlaGlyProProGlySerArgGlyAspProGlyPro 140

Qy       1829 GCTGGACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGA---AGGCAC 1773
                 |||         |||              ||||||||||||:::    |||
Db        141 ProGlyAlaProGlyPro---------AlaGlyProProGlySerArgGlyAspProGly 157

Qy       1772 TATCCAGGATGGCGAGGTCCAGGCAGATGCCCCGGC-----------CCGGAACCACCC 1725
                 ||||||      ||||||          ||||||            |||   ||||||
Db        158 ProProGlyAlaProGlyProAlaGlyProProGlySerArgGlyAspProGlyProPro 177

Qy       1724 TGG---CCTCGGTGGGCTCACCCACCACC-ACACGTACGGAGACATCACAGGCAGAGGCC 1669
                    |||       |||   ||||||    :::|||    :::                |||
Db        178 GlyAlaProGlyProAlaGlyProProGlySerArgGlyAspProGlyProProGlyAla 197

Qy       1668 CCGCAGAGCGCGGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTA 1609
                |||       ||||||     |||||||||      |||      |||      |||
Db        198 ProGlyProAlaGlyProProGlySerArgGlyAspProGlyProProGlyAlaProGly 217

Qy       1608 GGGAAGGGAGCTCCAGGCTTA------GGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCC 1555
                    |||   ||||||            ||||||
Db        218 ProAlaGlyProProGlySerArgGlyAspProGly------------------------ 229

Qy       1554 TCACTGCTAGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGG 1495
                          ||||||    :::||| |||       ||||||    |||    |||   ||
Db        230 ------------ProProGlyAlaPro-GlyPro---AlaGlyProProGlySerArgGl 244

Qy       1494 TGGTAGAGGGAGGCCAGTGTGT------AGGCAGGATCTGCAGGGCTGAGAAGGTGAAC 1441
               |    |||   |||                ||||||      |||    ||| |||:::|
Db        244 yAspProGlyProProGlyAlaProGlyProAlaGlyProProGlySerArg-GlyAspP 264

Qy       1440 CCGGTGAGGGCGGCTGAAGCTGTCACCACGGC-------------------CACACTG 1402
              |||||      |||            ||||||                        |||
Db        264 roGlyProProGly---------AlaHisGlyProAlaGlyProLysGlyAlaHis---- 279

Qy       1401 TGGGACAGGCATGTGGCACCGGCAGCCACAGGG-----------AAAGCTGCCACACTG 1354
                 |||   |||         |||:::|||  |||                 |||   ||| |||
Db        280 --GlyProAlaGlyProLysGlyAlaHis-GlyProAlaGlyProLysGlyAla-HisGl 298

Qy       1353 GCCAAATAGACTGCTCGAGTGCCGAATCGCTGCACCAGCCGGT---CCATGACCAGAGAG 1297
               |||                              |||   |||    |||    |||    |||
Db        298 yPro------------------------AlaGlyProLysGlyAlaProGlyProAlaGl 310

Qy       1296 AAGACCAGGGAGATGGCGCACTGCAGGAACAGCCCCAGGCTGCCCATCCGAACGCCTTCA 1237
                ||||||                   |||      ||||||  |||
Db        310 yProProGlySerArgGlyAspProGly---ProProGlyAlaPro-------------- 324

Qy       1236 TCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTACA----- 1182
                           ||||||     |||                  ||||||
Db        325 ------------GlyProAlaGlyProProGlySerArgGlyAspProGlyProProGl 340

Qy       1181 -GCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGCTG 1123
               ||||||  |.||    |||                                |||   |||
Db        340 yAlaProGlyProAlaGlyPro--------------------------ProGlySer-- 350

Qy       1122 CACAGCTCAGCCACGAAGAGCCGGCGCA------GGGTGCGGGGCATGCGGCAGCACAGC 1069
                      |||      ||||||        |||       ||||||
Db        351 -ArgGlyAspProGlyProProGlyAlaProGlyProAlaGlyProProGlySerArgGl 370

Qy       1068 T---------------GGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCC------------ 1035
                          ||||||    |||    ||||||     |||||||||
Db        370 yAspProGlyProProGlyAlaProGlyProAlaGlyProProGlySerArgGlyAspPr 390

Qy       1034 -GGAAAGCCAAGCGGGCCCGGCATGGACAGCAGTGGGGCGACAAGG------AGGGGGCC 982
               |||   |||      ||||||                     :::|||         |||||
```

```
Qy    231 AGCAGTTCTGGAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGT-- 288
Db    111 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro-GlnGlyLe 130

Qy    289 ----------CCAGAGGCTGTGGGTGAGCCG---CCTGCTGCGGCACCGGAAAGCCCAGC 335
Db    130 uProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGl 150

Qy    336 TCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCT 395
Db    150 yAlaProGlyThrProGlyGlu--------------------GlyGlnGlnHisHisLe 163

Qy    396 ATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGG 455
Db    163 uGlyGlyAlaArgGlnAlaGlyAspValGly----------------------------- 173

Qy    456 GCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACT 515
Db    174 -------SerProGlyAlaProGly---ThrProGlyProGln----------------- 184

Qy    516 GGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGC 575
Db    185 ---------------------------------------GlyLeuProGlySerProGl 191

Qy    576 TGAGCCT---CTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCT------GTGCC 626
Db    191 yAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPr 211

Qy    627 CGGATCCCAG-----------GCCCCTGGAGCT---GGCACTGCTCATCCTGGGCGTGG 671
Db    211 oGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr------ProGlyGluGl 229

Qy    672 G-----------GCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCC 719
Db    229 yGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaPr 249

Qy    720 TGCTCTCTGACCTCTTCCGGGACCC-----------------------GGACCACTGTC 755
Db    249 oGly---------ThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyTh 266

Qy    756 GCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCC 815
Db    266 rProGly-------------------------ProGlnGlyLeuProGlySerProGl 277

Qy    816 TGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCC------CTACCTGGGCACCCAGG 869
Db    277 yAla--------------ProGlyThrProGlyProGlnGlyLeuProGlySerPro-- 291

Qy    870 AGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGC 929
Db    292 -GlyAlaProGlyThrProGlyGluGlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAl 311

Qy    930 TGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCT 989
Db    311 aGlyAspValGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySe 331

Qy    990 CCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCG---GAACCTGGGCG 1046
Db    331 rProGlyAlaPro------GlyThrProGlyProGlnGlyLeuProGlySerProGlyAl 349

Qy   1047 CCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCT 1106
Db    349 aProGlyThrPro---GlyProGlnGlyLeuProGlySerPro---GlyAlaProGlyTh 367

Qy   1107 TCGTGGCTGAGCTGTGCAG------CTGGATGGCACTCATGACCTTCACGCTGTT----- 1155
Db    367 rProGlyGluGlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySe 387

Qy   1156 -------TTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGC 1208
Db    387 rProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla---Pr 406

Qy   1209 CGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCT-- 1266
Db    406 oGlyThrProGlyProGlnGlyLeu--------------ProGlySerProGlyAlaPr 421

Qy   1267 ----GTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGC 1322
Db    421 oGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGl 441

Qy   1323 AGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCG 1382
Db    441 uGlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAl 461

Qy   1383 GTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCA------ 1436
Db    461 aProGlyThrProGlyProGln---GlyLeuProGlySerProGlyAlaProGlyThrPr 480

Qy   1437 ----------------CCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTG 1480
Db    480 oGlyProGlnGlyLeuProGlySerProGly-AlaProGlyThrProGly---ProGlnG 499

Qy   1481 GCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCT------------------------ 1515
Db    499 lyLeuProGlySerProGlyAlaProGlyThrProGlyGluGlyGlnGlnHisHisLeuG 519

Qy   1516 -----GCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACC 1570
Db    519 lyGlyAla-ArgGlnAlaGlyAspValGlySerProGlyAlaProGlyThrProGlyPro 538

Qy   1571 AGCTTCCTGCCAGGCCCT------AAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGT 1624
Db    539 GlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGly 558

Qy   1625 GCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCC------------ 1672
Db    559 SerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro 578

Qy   1673 ---------------------------------------------TCTGCCTGTGATGTC 1687
Db    579 GlyThrProGlyGluGlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAsp--- 597

Qy   1688 TCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATC 1747
Db    598 ---------------ValGly-SerProGlyAlaPro-GlyThrProGlyProGlnGlyL 612

Qy   1748 TGCCTGGA------------CCTCGCCATCCTGGA---TAGTGCCTTCCTGCTGTCCCAG 1792
Db    612 euProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProG 632

Qy   1793 GTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTAT 1852
Db    632 lyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro---------------- 646

Qy   1853 ATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTT 1912
Db    647 --GlyAlaProGlyThrProGlyGluGlyGln---------------------------- 656

Qy   1913 GACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG---- 1968
Db    657 --GlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaProG 676

Qy   1969 --------------GAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCC 2014
Db    676 lyThrProGlyProGlnGlyLeuProGlySerProGlyAla-ProGlyThrProGlyPro 695

Qy   2015 ATGGGGCTGCCGGGCTGGCCG 2035
Db    696 GlnGlyLeuProGlySerPro 702
```

RESULT 31
US-08-642-255-62
; Sequence 62, Application US/08642255
; Patent No. 5773249
; GENERAL INFORMATION:

```
                   ||||||            |||||||||  |||   |||     ||||||
Db        331 rProGlyAlaPro------GlyThrProGlyProGlnGlyLeuProGlySerProGlyAl 349

Qy       1047 CCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCT 1106
                |||   :::|||      |||     ||||||   :::|||      ||||||
Db        349 aProGlyThrPro---GlyProGlnGlyLeuProGlySerPro---GlyAlaProGlyTh 367

Qy       1107 TCGTGGCTGAGCTGTGCAG------CTGGATGGCACTCATGACCTTCACGCTGTT----- 1155
                    |||         |||      |||   |||                |||
Db        367 rProGlyGluGlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySe 387

Qy       1156 -------TTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGC 1208
                             |||     |||       :::||||||         :::
Db        387 rProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla---Pr 406

Qy       1209 CGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCT-- 1266
                |||      ||||||:::   |||                  |||   |||||||||
Db        406 oGlyThrProGlyProGlnGlyLeu--------------ProGlySerProGlyAlaPr 421

Qy       1267 ----GTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGC 1322
                   |||       :::   |||||||||                |||       |||
Db        421 oGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGl 441

Qy       1323 AGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCG 1382
                          :::|||                   |||    |||
Db        441 uGlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAl 461

Qy       1383 GTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCA------ 1436
                          |||    ||||||    |||          |||        |||
Db        461 aProGlyThrProGlyProGln---GlyLeuProGlySerProGlyAlaProGlyThrPr 480

Qy       1437 ---------------CCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTG 1480
                                |||||||||||     :::|||        |||           |
Db        480 oGlyProGlnGlyLeuProGlySerProGly-AlaProGlyThrProGly---ProGlnG 499

Qy       1481 GCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCT------------------------ 1515
              ||||||||        ||||||      |||   |||
Db        499 lyLeuProGlySerProGlyAlaProGlyThrProGlyGluGlyGlnGlnHisHisLeuG 519

Qy       1516 -----GCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACC 1570
                   ||| :::         ||||||    |||          :::      :::
Db        519 lyGlyAla-ArgGlnAlaGlyAspValGlySerProGlyAlaProGlyThrProGlyPro 538

Qy       1571 AGCTTCCTGCCAGGCCCCT------AAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGT 1624
                     |||||||||            ||||||   |||    |||    |||      |||
Db        539 GlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGly 558

Qy       1625 GCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCC------------ 1672
              :::    |||:::          |||    |||        |||    |||:::
Db        559 SerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro 578

Qy       1673 --------------------------------------------TCTGCCTGTGATGTC 1687
                                                           |||    |||
Db        579 GlyThrProGlyGluGlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAsp--- 597

Qy       1688 TCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATC 1747
                                ||||||  ||||||        |||  |||      ||||||        |
Db        598 --------------ValGly-SerProGlyAlaPro-GlyThrProGlyProGlnGlyL 612

Qy       1748 TGCCTGGA-----------CCTCGCCATCCTGGA---TAGTGCCTTCCTGCTGTCCCAG 1792
              ||||||||              |||     ||||||           ||||||         ||||
Db        612 euProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProG 632

Qy       1793 GTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTAT 1852
               ||   |||                                |||       |||
Db        632 lyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro--------------- 646

Qy       1853 ATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTT 1912
                 |||            ||||||   |||:::
Db        647 --GlyAlaProGlyThrProGlyGluGlyGln--------------------------- 656

Qy       1913 GACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG---- 1968
                 |||        ||||||                        :::          |||
Db        657 --GlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaProG 676

Qy       1969 -------------GAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCC 2014
                           :::|||||||||             ||| |||            |||
Db        676 lyThrProGlyProGlnGlyLeuProGlySerProGlyAla-ProGlyThrProGlyPro 695

Qy       2015 ATGGGGCTGCCGGGCTGGCCG 2035
                 ||||||||||||   |||
Db        696 GlnGlyLeuProGlySerPro 702
```

```
RESULT 30
US-08-397-633A-31
; Sequence 31, Application US/08397633A
; Patent No. 5773577
;  GENERAL INFORMATION:
;    APPLICANT:  Cappello, Joseph
;    TITLE OF INVENTION:  PRODUCTS COMPRISING SUBSTRATESCAPABLE
;    TITLE OF INVENTION:  OF ENZYMATIC CROSS-LINKING
;    NUMBER OF SEQUENCES:  105
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/397,633A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rowland, Bertram I
;      REGISTRATION NUMBER:  20,015
;      REFERENCE/DOCKET NUMBER:  A-58848-1/BIR PROP-011-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 781-1989
;      TELEFAX:  (415) 398-3249
;      TELEX:  910 277299
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  762 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-397-633A-31

Alignment Scores:
Pred. No.:              6.7e-20        Length:          762
Score:                  364.00         Matches:         223
Percent Similarity:     33.95%         Conservative:    32
Best Local Similarity:  29.69%         Mismatches:      300
Query Match:            5.67%          Indels:          197
DB:                     1              Gaps:            43

US-09-759-143-110 (1-3410) x US-08-397-633A-31 (1-762)

Qy         84 GGTGCCCCACAGCAGCAGGTGTTG------AGCATGGGCTGAGAAGCTGGACCGGCACCA 137
                 |||   |||      |||   |||       |||   |||            |||   |||
Db         52 GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro 71

Qy        138 AAGGGC------TGGCAGAAATGGGCGCCTGGC--------------------TGATTC 170
                 :::|||                 |||||||||
Db         72 GlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGluGlyGlnGlnHisHis 91

Qy        171 CTAGGCAGTTGGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGA 230
                ||||||      |||   |||   |||:::          :::   ||||||
Db         92 LeuGlyGlyAlaArg---GlnAlaGlyAspValGlySerProGlyAlaProGlyThrPro 110
```

```
Qy      269 GTAGGGCTCAGGGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCA-- 212
                |||||| |||     ||| |||                        ||||||
Db      784 uLysGlySerProGlyAlaGlnGlyProPro-----------------GlyAlaProGl 798

Qy      211 ------------------------------GAAGCTGCGGCCTCTCCTCCTTGCTGCCG 183
                                         |||   ||||||
Db      798 yProLeuGlyIleAlaGlyIleThrGlyAlaArgGlyLeuAlaGlyProProGlyMetPr 818

Qy      182 CCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGC 123
                    ||||||    ||||||           :::             |||
Db      818 oGly---ProArgGlySerProGlyProGlnGlyValLysGlyGluSerGlyLysProGl 837

Qy      122 TTCTCAG-----------------------------CCCATGCTCAACACCTGCTGCTG 93
             :::                                |||         |||
Db      837 yAlaAsnGlyLeuSerGlyGluArgGlyProProGlyProGln-GlyLeuProGlyLeuA 857

Qy       92 TGGGGCACCTCAGTGGGGACACGTCTCATCACTCAGATCCTGGC---------------- 49
              |||     :::|||:::               :::||||||
Db      857 laGly---ThrAlaGlyGluProGlyArgAspGlyAsnProGlySerAspGlyLeuProG 876

Qy       48 -------------------------------CGAGGCGCGCGGCTGTCACCCGGAGCC 22
                                        ||||||      |||||||||||
Db      876 lyArgAspGlySerProGlyGlyLysGlyAspArgGlyGluAsnGlySerProGlyAla 895
```

RESULT 29
US-08-642-255-120
; Sequence 120, Application US/08642255
; Patent No. 5773249
; GENERAL INFORMATION:
; APPLICANT: CAPPELLO, Joseph
; APPLICANT: FERRARI, Franco A.
; TITLE OF INVENTION: High Molecular Weight Collagen-Like
; TITLE OF INVENTION: Protein Polymers
; NUMBER OF SEQUENCES: 135
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/642,255
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ROWLAND, Bertram I.
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A55556-3/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 120:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 762 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-642-255-120

Alignment Scores:

Pred. No.:	6.7e-20	Length:	762
Score:	364.00	Matches:	223
Percent Similarity:	33.95%	Conservative:	32
Best Local Similarity:	29.69%	Mismatches:	300
Query Match:	5.67%	Indels:	197
DB:	1	Gaps:	43

US-09-759-143-110 (1-3410) x US-08-642-255-120 (1-762)

```
Qy       84 GGTGCCCCACAGCAGCAGGTGTTG------AGCATGGGCTGAGAAGCTGGACCGGCACCA 137
            |||  |||     |||  |||          |||  |||          |||   |||
Db       52 GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro 71

Qy      138 AAGGGC------TGGCAGAAATGGGCGCCTGGC--------------------TGATTC 170
            :::|||                |||||||||
Db       72 GlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGluGlyGlnGlnHisHis 91

Qy      171 CTAGGCAGTTGGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGA 230
            ||||||      |||   |||   |||:::      :::   ||||||
Db       92 LeuGlyGlyAlaArg---GlnAlaGlyAspValGlySerProGlyAlaProGlyThrPro 110

Qy      231 AGCAGTTCTGGAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGT-- 288
                      |||:::|||      |||      |||      |||||||||   |||
Db      111 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro-GlnGlyLe 130

Qy      289 ---------CCAGAGGCTGTGGGTGAGCCG---CCTGCTGCGGCACCGGAAAGCCCAGC 335
                     |||   |||   |||   |||   |||          |||   ||||||
Db      130 uProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGl 150

Qy      336 TCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCT 395
                  |||   |||                                |||::::::||||||||
Db      150 yAlaProGlyThrProGlyGlu--------------------GlyGlnGlnHisHisLe 163

Qy      396 ATGTGCCGCCTCTGCTGCTGGAAGTGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGG 455
             |      |||       ||||||      |||
Db      163 uGlyGlyAlaArgGlnAlaGlyAspValGly---------------------------- 173

Qy      456 GCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACT 515
                      |||       ||||||      |||    |||:::
Db      174 -------SerProGlyAlaProGly---ThrProGlyProGln---------------- 184

Qy      516 GGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGC 575
                                                          :::    |||   |||
Db      185 ---------------------------------------GlyLeuProGlySerProGl 191

Qy      576 TGAGCCT---CTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCT------GTGCC 626
                  |||          |||:::|||        |||       |||          ||
Db      191 yAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPr 211

Qy      627 CGGATCCCAG-----------GCCCCTGGAGCT---GGCACTGCTCATCCTGGCGTGG 671
            ||||  |||              :::|||||||||   ||||||        ||||||   ||
Db      211 oGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr------ProGlyGluGl 229

Qy      672 G-----------GCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCC 719
            |                  |||             |||   :::   |||   |||   ||
Db      229 yGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaPr 249

Qy      720 TGCTCTCTGACCTCTTCCGGGACCC----------------------GGACCACTGTC 755
            |                  |||||||||                          |||   ::
Db      249 oGly---------ThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyTh 266

Qy      756 GCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCC 815
            :||||||                                |||||||||||   |||
Db      266 rProGly--------------------------ProGlnGlyLeuProGlySerProGl 277

Qy      816 TGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCC------CTACCTGGGCACCCAGG 869
             |||                     |||||||||      ||||||||   |||
Db      277 yAla--------------ProGlyThrProGlyProGlnGlyLeuProGlySerPro-- 291

Qy      870 AGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGC 929
             |||   |||         |||            ||||||         :::      ||
Db      292 -GlyAlaProGlyThrProGlyGluGlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAl 311

Qy      930 TGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCT 989
             ||||     |||            |||                   :::   :::
Db      311 aGlyAspValGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySe 331

Qy      990 CCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCG---GAACCTGGGCG 1046
```

```
Db     214 ------------------------SerAsnGlyAlaPro-----------GlyGlnAr 221

Qy    2129 AGGGACGCCCCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAG 2070
           ||||  |||  |||                        :::  ||||||||
Db     221 gGlyGluProGlyProGlnGly-----------HisAlaGlyAlaGlnGlyProProGl 237

Qy    2069 AGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGC 2010
              |||         ||||||                          |||  |||
Db     237 yProProGlyIleAsnGlySer--------------------------ProGlyGly-- 247

Qy    2009 TAACAGGAGCGGGGAGCTGGGACCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAG 1950
                  ||||||:::|||||||:::  |||  |||      |||
Db     248 -------LysGlyGluMetGlyProAla--GlyIleProGlyAlaPro----------- 260

Qy    1949 TTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGT 1890
                       :::     :::|||          |||          :::
Db     261 -----------GlyLeuMetGlyAlaArgGly-----------ProProGlyProAla 272

Qy    1889 AAATGGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGA 1830
                          ||||||||||||                         |||
Db     273 GlyAlaAsnGlyAlaProGlyLeuArg-------------------------GlyGly 283

Qy    1829 GCTGGACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCAC--- 1773
           ||||||:::                       |||   ||||||    |||
Db     284 AlaGlyGluProGlyLysAsnGlyAlaLysGlyGluProGlyProArgGlyGluArgGly 303

Qy    1772 -----------------TATCCAGGATGGCGAGGT--------------------CCA 1752
                            ||||||   :::|||                       |||
Db     304 GluAlaGlyIleProGlyValProGlyAlaLysGlyGluAspGlyLysAspGlySerPro 323

Qy    1751 GGCAGATGCCCCGGCCCGGAACCACCCTGGCCTCGGTGGGCTCACCCACCACCACACGTA 1692
           ||| |||    |||       ||||||   |||
Db     324 Gly-AspProGlyAlaAsnGlyLeuProGlyAlaAlaGly--------------GluAr 338

Qy    1691 CGGAGACATCACAGGCAGAGGCCCCGCA--------------------GAGCGCGGGTG 1653
            |||   :::      ||||||||||||                   |||::: |||
Db     338 gGlyAlaLeuGlySerArgGlyProAlaGlyProAsnGlyIleProGlyGluLys-GlyP 358

Qy    1652 GAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGGG---AAGGGAGCTC 1596
              |||  |||        |||  |||   |||  |||        |||  |
Db     358 roAlaGlyGluArgGly---AlaProGlyProAlaGlyProArgGlyAlaAlaGlyGluP 377

Qy    1595 CAGGCTTAGGG------CCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCTAGCAC 1542
           |||||            |||||| |||        |||  |||
Db     377 roGlyArgAspGlyValProGly-GlyProGlyMetArgGlyMetPro----------- 392

Qy    1541 CTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGTAGAGGGAGG 1482
                                   |||:::|||     ||||||    |||:::
Db     393 -------------------------GlySerProGlyGlyProGlySerAspGlyLys 403

Qy    1481 CCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAG 1422
           |||        |||||| ||||||      |||   ||||||     |||   ||
Db     404 Pro---GlyProProGlySer-GlnGlyGluSerGlyArgProGlyProProGlyProSe 422

Qy    1421 CTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACCGGCAGCCACAGGGAAAGCTG 1362
           |   ||||||        |||     |||        |||:::|||
Db     422 r--GlyProArgGlyGlnProGlyVal---MetGlyPheProGlyProLysGlyAsn--- 439

Qy    1361 CCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCTGCACCAGCCGGTCCATGACCA 1302
                 |||   :::   |||            |||  ||||||   |||
Db     440 AspGlyAlaProGlyLysAsnGlyGlu------ArgGlyGlyProGlyGlyProGlyPro 457

Qy    1301 GAGAGAAGACCAGGGAGATGGCGCACTGCAGGA--------------------ACAGCC 1263
           :::       ||||||:::      |||                          |||
Db     458 GlnGlyProProGlyLysAsnGlyGluTyrGlyProGlnGlyProProGlyProThrGly 477

Qy    1262 CCAGGCTGCCCATCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAG 1203
           ||||||                            :::|||||||||   |||
Db     478 ProGlyGlyAspLys---------------GlyAspThrGlyProArgGlyProGlnGly 492

Qy    1202 CTCTGGGCACGCCCTGGTACA------------------------------------ 1182
           |||        |||||||||
Db     493 LeuGlnGlyLeuProGlyThrGlyGlyProProGlyGluAsnGluLysProGlyGluPro 512

Qy    1181 ---------------------------------------------------------GCC 1179
                                                                     |||
Db     513 GlyProLysGlyGluAlaGlyAlaProGlyAlaProGlyGlyLysGlyAspAlaGlyAla 532

Qy    1178 CCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGCTGCACA 1119
           |||    |||   |||                                   |||
Db     533 ProGlyGluArgGlyPro-------------------------ProGlyLeuAlaGly 543

Qy    1118 GCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACAGCTGGTGCAGCC 1059
           |||       |||  |||||||||                             |||||||||
Db     544 AlaProGlyLeuArgGlyGlyAlaGlyProProGlyProGluGlyGlyLysGlyAlaAla 563

Qy    1058 GGGGAAGCAGGGCGCCCAGGTTCC---------------------------------- 1035
           |||      |||   ||||||:::
Db     564 GlyProProGlyProProGlyAlaAlaGlyThrProGlyLeuGlnGlyMetProGlyGlu 583

Qy    1034 -------------------------------------GGAAAGCCAAGCGGGCCCGGC 1014
                                                |||:::|||  |||||||||
Db     584 ArgGlyGlyLeuGlySerProGlyProLysGlyAspLysGlyGluProGlyGlyProGly 603

Qy    1013 ATGGACAGCAGTGGGGCGACA------AGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGG 960
              |||                       ||||||||||||
Db     604 AlaAspGlyValProGlyLysAspGlyProArgGlyProThr----------------- 617

Qy     959 TGGGGCCCAGCGCTGCCTCCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGA 900
              ||||||     ||||||   |||                          |||:::
Db     618 ---GlyProIleGlyProProGlyProAla----------------------GlyGln 628

Qy     899 TGAGGGTGAGCAGGCCAAAGAGGCACTCCTCCTGGGTGCCCAGGT---AGGGGGCCAGGG 843
           |||      |||:::      ||||||   ||||||  ||||||
Db     629 ProGlyAspLysGlyGluGlyGlyAla---ProGlyLeuProGlyIleAlaGlyPro--- 646

Qy     842 CACTGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGACTGATCA 783
                                    ||||||||||||
Db     647 -------------------------ArgGlySerProGlyGlu-------------- 652

Qy     782 TGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGT---CCCGGAAGAGGT 729
              |||              ||||||  :::|||  ||||||   ||||||:::
Db     653 ---ArgGlyGluThrGlyProProGlyProAlaGlyPheProGlyAlaProGlyGlnAsn 671

Qy     728 CAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCA 669
                  |||       |||         ||||||   :::              |||
Db     672 GlyGluProGlyGlyLysGlyGluArgGlyAlaProGlyGluLysGlyGluGlyGlyPro 691

Qy     668 CGCCCAGGATGAGCAGTGCCAGCTCCAGGG---------GCCTGGGATCCGGGCACAGCA 618
              |||||| |||||||||     |||         |||      ||||||
Db     692 ---ProGlyValAlaValProProGlyGlySerGlyProAlaGlyProProGlyProGln 710

Qy     617 GCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCAAGGACA 558
              :::                 ||||||||||||   |||   |||
Db     711 GlyValLysGly-----------------GluArgGlySerProGlyGlyPro------ 722

Qy     557 GTGCCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAGC 498
                                                   |||:::|||
Db     723 ----------------------------------GlyAlaAlaGlyPhe--------- 727

Qy     497 CTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCATGGTCATGA 438
              ||||||     |||||||||                    |||   |||
Db     728 ProGlyAlaArgGlyLeuProGlyPro--------------------ProGlySerAsn 740

Qy     437 ACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCCA 378
              :::|||  ||||||  |||:::
Db     741 GlyAsnProGlyProProGlyProSerGlySer------------------------- 751

Qy     377 AACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGTGCC 318
                    ||||||:::      |||   |||     ||||||   :::||||||
Db     752 ---------ProGlyLys------AspGlyProProGlyProAlaGlyAsnThrGlyAla 766

Qy     317 GCAGCAGGCGGCTCACCCACAGCCTCTGGACCATAGTGGG-----------CCAGGCGG 270
                 |||||||||   |||||||||   |||           ||||||
Db     767 Pro------GlySerProGlyValSerGlyPro-LysGlyAspAlaGlyGlnProGlyGl 784
```

```
                    ||||||:::            |||  |||       ||||||   :::||||||
Db      752 ---------ProGlyLys------AspGlyProProGlyProAlaGlyAsnThrGlyAla 766

Qy      317 GCAGCAGGCGGCTCACCCACAGCCTCTGGACCATAGTGGG-----------CCAGGCGG 270
                    |||||||||     |||||||||    |||          ||||||
Db      767 Pro------GlySerProGlyValSerGlyPro-LysGlyAspAlaGlyGlnProGlyGl 784

Qy      269 GTAGGGCTCAGGGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCA-- 212
                ||||||   |||    |||   |||                     ||||||
Db      784 uLysGlySerProGlyAlaGlnGlyProPro-----------------GlyAlaProGl 798

Qy      211 ------------------------------GAAGCTGCGGCCTCTCCTCCTTGCTGCCG 183
                                               |||    ||||||
Db      798 yProLeuGlyIleAlaGlyIleThrGlyAlaArgGlyLeuAlaGlyProProGlyMetPr 818

Qy      182 CCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGC 123
                  ||||||        ||||||          :::           |||
Db      818 oGly---ProArgGlySerProGlyProGlnGlyValLysGlyGluSerGlyLysProGl 837

Qy      122 TTCTCAG----------------------------CCCATGCTCAACACCTGCTGCTG 93
             :::                                |||        |||
Db      837 yAlaAsnGlyLeuSerGlyGluArgGlyProProGlyProGln-GlyLeuProGlyLeuA 857

Qy       92 TGGGGCACCTCAGTGGGGACACGTCTCATCACTCAGATCCTGGC--------------- 49
              |||         :::|||:::              :::||||||
Db      857 laGly---ThrAlaGlyGluProGlyArgAspGlyAsnProGlySerAspGlyLeuProG 876

Qy       48 ------------------------------CGAGGCGCGCGGCTGTCACCCGGAGCC 22
                                            ||||||         ||||||||||||
Db      876 lyArgAspGlySerProGlyGlyLysGlyAspArgGlyGluAsnGlySerProGlyAla 895
```

RESULT 28
US-09-548-608-21
; Sequence 21, Application US/09548608
; Patent No. 6355442
; GENERAL INFORMATION:
; APPLICANT: Qvist, Per
; APPLICANT: Bonde, Martin
; TITLE OF INVENTION: A Method for Assaying Collagen Fragments
; TITLE OF INVENTION: in Body Fluids, A Test Kit and Means for Carrying Out the
; TITLE OF INVENTION: Method and Use of the Method to Diagnose the Presence of
; TITLE OF INVENTION: Disorders Associated with the Metabolism of
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/548,608
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/187,319
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Gogoris, Adda C
; REGISTRATION NUMBER: 29,714
; REFERENCE/DOCKET NUMBER: 4305/08701
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-527-7700
; TELEFAX: 212-753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1078 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: COLLAGEN ALPHA 1 (III)
US-09-548-608-21

Alignment Scores:

Pred. No.:	6.94e-20	Length:	1078
Score:	364.50	Matches:	290
Percent Similarity:	30.79%	Conservative:	57
Best Local Similarity:	25.73%	Mismatches:	354
Query Match:	5.86%	Indels:	426
DB:	4	Gaps:	62

US-09-759-143-110 (1-3410) x US-09-548-608-21 (1-1078)

```
Qy     2939 GGAAGTGGGGGGAACCAGGCTGGGCCAAGAGAAGAGGGGTGGTTAGGGAAGCCGTTGAGA 2880
              ||||||   |||   |||      |||   |||   |||      |||   |||
Db       48 GlySerProGlyTyrGlnGlyProProGlyGluProGlyGlnAlaGly---ProSerGly 66

Qy     2879 CCTGAAGCCCCACCCTCTACCTTCCTTCAACACCCTAACCTTGGGTAACAGCATTTGGAA 2820
             |||   ||||||
Db       67 ProProGlyProProGly------------------------------------------ 72

Qy     2819 TTATCATTTGGGATGAGTAGAATTTCCAAGGTCCTGGGTTAGGCATTTTGGGGGGCCAGA 2760
               :::   |||   |||      :::             |||   :::         |||
Db       73 ---AlaIleGlyProSerGlyProAlaGlyLysAspGlyGluSerGlyArgProGlyArg 91

Qy     2759 CCCCAGGAGAAGAAGATTCTGGCAATGATCAGCCCAATGACCAGCTATCTCAGGGGACCT 2700
             |||   ::::::   :::                  |||          ::::::||||||
Db       92 ProGlyAspArgGlyLeu--------------ProGlyProProGlyIleLysGlyPro 106

Qy     2699 GATTGTTGGGGATCCCCCACCCTACCCAAATATTAGACACCAACACAGA---------- 2651
                     |||   |||   :::|||   ::::::      ||||||
Db      107 Ala------GlyIle-ProGlyPheProGlyMetLys---GlyHisArgGlyPheAspGl 123

Qy     2650 ------------AAAGCTAGCAATGGATTCCCTTCTACTTTGTTAAATAAATAAGTTAA 2604
                          |||         |||   |||                      :::||
Db      123 yArgAsnGlyGluLysGlyGluThrGlyAlaProGly-----------------LeuLy 137

Qy     2603 ATATTTAAATGCCTGTGTCTCTGTGATGGCAACAGAAGGACCAACAGGCCACATCCTGAT 2544
              |      |||                           ||||||   |||
Db      137 sGlyGluAsnGlyLeuProGlyGluAsnGlyAlaProGlyProMetGlyProArgGlyAl 157

Qy     2543 AAAAGGTAAGAGGGGGGTGGATCAGCAAAAAGACAGTGCTGTGGGCTGAGGGGACCTGGT 2484
                  |||:::||||||                     :::|||      |||:::
Db      157 aProGlyGluArgGlyArgProGlyLeuProGlyAlaAlaGlyAlaArgGlyAsn----- 175

Qy     2483 TCTTGTGTGTTGCCCCTCAGGACTCTTCCCCTACAAATAAGTCATATGTTCAAATCCCAT 2424

Db      175 ------------------------------------------------------------ 175

Qy     2423 GGAGGAGTGTTTCATCCTAGAAACTCCCATGCAAGAGCTACATTAAACGAAGCTGCAGGT 2364

Db      175 ------------------------------------------------------------ 175

Qy     2363 TAAGGGGCTTAGA------GATGGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAA 2310
                |||   |||      ||||||:::||||||                  |||
Db      176 -AspGlyAlaArgGlySerAspGlyGlnProGly-----------------ProProGl 189

Qy     2309 CCCTTCTCTAGGTGTGTCTCAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCC 2250
              |||      |||   :::                                 |||:::||
Db      189 yProPro---GlyThrAlaGlyPhePro-----------------------GlySerPr 200

Qy     2249 ACCTGCAGAGTCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCA 2190
             |   |||:::                  |||   ||||||
Db      200 oGlyAlaLysGlyGluValGlyProAlaGlySerProGly------------------- 213

Qy     2189 GACTGAAACCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAG 2130
                                       :::|||:::|||                :::||
```

```
Qy      2249 ACCTGCAGAGTCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCA 2190
             |  |||:::                        |||   ||||||
Db       200 oGlyAlaLysGlyGluValGlyProAlaGlySerProGly------------------- 213

Qy      2189 GACTGAAACCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAG 2130
                                     :::|||:::|||             :::||
Db       214 -----------------------SerAsnGlyAlaPro-----------GlyGlnAr 221

Qy      2129 AGGGACGCCCCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAG 2070
             ||||  |||  |||                        :::  ||||||||
Db       221 gGlyGluProGlyProGlnGly-----------HisAlaGlyAlaGlnGlyProProGl 237

Qy      2069 AGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGC 2010
                |||      ||||||                          |||  |||
Db       237 yProProGlyIleAsnGlySer--------------------------ProGlyGly-- 247

Qy      2009 TAACAGGAGCGGGGAGCTGGGACCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAG 1950
                     ||||||:::||||||:::  |||   |||        |||
Db       248 -------LysGlyGluMetGlyProAla--GlyIleProGlyAlaPro----------- 260

Qy      1949 TTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGT 1890
                         :::     :::|||           |||        :::
Db       261 -----------GlyLeuMetGlyAlaArgGly-----------ProProGlyProAla 272

Qy      1889 AAATGGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGA 1830
                               |||||||||||                         |||
Db       273 GlyAlaAsnGlyAlaProGlyLeuArg-------------------------GlyGly 283

Qy      1829 GCTGGACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCAC--- 1773
             ||||||:::                    |:|  ||||||     |||
Db       284 AlaGlyGluProGlyLysAsnGlyAlaLysGlyGluProGlyProArgGlyGluArgGly 303

Qy      1772 ------------------TATCCAGGATGGCGAGGT--------------------CCA 1752
                                |||||   :::|||                     |||
Db       304 GluAlaGlyIleProGlyValProGlyAlaLysGlyGluAspGlyLysAspGlySerPro 323

Qy      1751 GGCAGATGCCCCGGCCCGGAACCACCCTGGCCTCGGTGGGCTCACCCACCACCACACGTA 1692
             ||| |||      |||      ||||||     |||
Db       324 Gly-AspProGlyAlaAsnGlyLeuProGlyAlaAlaGly--------------GluAr 338

Qy      1691 CGGAGACATCACAGGCAGAGGCCCCGCA--------------------GAGCGCGGGTG 1653
              |||   :::      |||||||||||                     |||::: |||
Db       338 gGlyAlaLeuGlySerArgGlyProAlaGlyProAsnGlyIleProGlyGluLys-GlyP 358

Qy      1652 GAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGGG---AAGGGAGCTC 1596
               |||  |||       |||  |||   |||   |||      |||    |
Db       358 roAlaGlyGluArgGly---AlaProGlyProAlaGlyProArgGlyAlaAlaGlyGluP 377

Qy      1595 CAGGCTTAGGG------CCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCTAGCAC 1542
             |||||             |||||| |||          |||   |||
Db       377 roGlyArgAspGlyValProGly-GlyProGlyMetArgGlyMetPro----------- 392

Qy      1541 CTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGTAGAGGGAGG 1482
                                        |||:::|||       ||||||      |||:::
Db       393 --------------------------GlySerProGlyGlyProGlySerAspGlyLys 403

Qy      1481 CCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAG 1422
             |||            |||||| ||||||      |||    ||||||      |||  ||
Db       404 Pro---GlyProProGlySer-GlnGlyGluSerGlyArgProGlyProProGlyProSe 422

Qy      1421 CTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACCGGCAGCCACAGGGAAAGCTG 1362
             |   ||||||         |||         |||            |||:::|||
Db       422 r--GlyProArgGlyGlnProGlyVal---MetGlyPheProGlyProLysGlyAsn--- 439

Qy      1361 CCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCTGCACCAGCCGGTCCATGACCA 1302
                      |||   :::    |||               |||  ||||||  |||
Db       440 AspGlyAlaProGlyLysAsnGlyGlu------ArgGlyGlyProGlyGlyProGlyPro 457

Qy      1301 GAGAGAAGACCAGGGAGATGGCGCACTGCAGGA--------------------ACAGCC 1263
             :::        ||||||:::          |||                    |||
Db       458 GlnGlyProProGlyLysAsnGlyGluTyrGlyProGlnGlyProProGlyProThrGly 477

Qy      1262 CCAGGCTGCCCATCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAG 1203
             ||||||                              :::|||||||||    |||
Db       478 ProGlyGlyAspLys--------------GlyAspThrGlyProArgGlyProGlnGly 492

Qy      1202 CTCTGGGCACGCCCTGGTACA--------------------------------------- 1182
             |||          |||||||||
Db       493 LeuGlnGlyLeuProGlyThrGlyGlyProProGlyGluAsnGluLysProGlyGluPro 512

Qy      1181 ---------------------------------------------------------GCC 1179
                                                                      |||
Db       513 GlyProLysGlyGluAlaGlyAlaProGlyAlaProGlyGlyLysGlyAspAlaGlyAla 532

Qy      1178 CCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGCTGCACA 1119
             |||      |||  |||                             |||
Db       533 ProGlyGluArgGlyPro-------------------------ProGlyLeuAlaGly 543

Qy      1118 GCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACAGCTGGTGCAGCC 1059
             |||            |||·  |||||||||                       |||||||||
Db       544 AlaProGlyLeuArgGlyGlyAlaGlyProProGlyProGluGlyGlyLysGlyAlaAla 563

Qy      1058 GGGGAAGCAGGGCGCCCAGGTTCC------------------------------------ 1035
             |||      |||   |||||||:::
Db       564 GlyProProGlyProProGlyAlaAlaGlyThrProGlyLeuGlnGlyMetProGlyGlu 583

Qy      1034 -------------------------------------GGAAAGCCAAGCGGGCCCGGC 1014
                                                  |||:::|||  |||||||||
Db       584 ArgGlyGlyLeuGlySerProGlyProLysGlyAspLysGlyGluProGlyGlyProGly 603

Qy      1013 ATGGACAGCAGTGGGGCGACA------AGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGG 960
                |||                        |||||||||||
Db       604 AlaAspGlyValProGlyLysAspGlyProArgGlyProThr----------------- 617

Qy       959 TGGGGCCCAGCGCTGCCTCCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGA 900
                |||||     ||||||  |||                          |||:::
Db       618 ---GlyProIleGlyProProGlyProAla-----------------------GlyGln 628

Qy       899 TGAGGGTGAGCAGGCCAAAGAGGCACTCCTCCTGGGTGCCCAGGT---AGGGGGCCAGGG 843
                |||   |||:::   |||      ||||||   ||||||     ||||||
Db       629 ProGlyAspLysGlyGluGlyGlyAla---ProGlyLeuProGlyIleAlaGlyPro--- 646

Qy       842 CACTGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGACTGATCA 783
                                         |||||||||||
Db       647 -------------------------ArgGlySerProGlyGlu--------------- 652

Qy       782 TGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGT---CCCGGAAGAGGT 729
                |||                   ||||||   :::|||   ||||||   ||||||:::
Db       653 ---ArgGlyGluThrGlyProProGlyProAlaGlyPheProGlyAlaProGlyGlnAsn 671

Qy       728 CAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCA 669
                   |||      |||      ||||||   :::                   |||
Db       672 GlyGluProGlyGlyLysGlyGluArgGlyAlaProGlyGluLysGlyGluGlyGlyPro 691

Qy       668 CGCCCAGGATGAGCAGTGCCAGCTCCAGGG---------GCCTGGGATCCGGGCACAGCA 618
                ||||||  |||||||||      |||        |||      ||||||
Db       692 ---ProGlyValAlaValProProGlyGlySerGlyProAlaGlyProProGlyProGln 710

Qy       617 GCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCAAGGACA 558
                :::                               |||||||||||  |||  |||
Db       711 GlyValLysGly-----------------GluArgGlySerProGlyGlyPro------ 722

Qy       557 GTGCCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAGC 498
                                                  |||:::|||
Db       723 -----------------------------------GlyAlaAlaGlyPhe--------- 727

Qy       497 CTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCATGGTCATGA 438
                ||||||   |||||||||                          |||   |||
Db       728 ProGlyAlaArgGlyLeuProGlyPro-------------------ProGlySerAsn 740

Qy       437 ACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCCA 378
                :::|||   ||||||   |||:::
Db       741 GlyAsnProGlyProProGlyProSerGlySer------------------------- 751

Qy       377 AACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGTGCC 318
```

```
Db        728 ProGlyAlaArgGlyLeuProGlyPro--------------------ProGlySerAsn 740

Qy        437 ACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCCA 378
                 :::|||   ||||||   |||:::
Db        741 GlyAsnProGlyProProGlyProSerGlySer-------------------------- 751

Qy        377 AACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGTGCC 318
                      ||||||:::      |||  |||      ||||||  :::||||||
Db        752 ---------ProGlyLys------AspGlyProProGlyProAlaGlyAsnThrGlyAla 766

Qy        317 GCAGCAGGCGGCTCACCCACAGCCTCTGGACCATAGTGGG-----------CCAGGCGG 270
                 |||||||||    |||||||||  |||             ||||||
Db        767 Pro------GlySerProGlyValSerGlyPro-LysGlyAspAlaGlyGlnProGlyGl 784

Qy        269 GTAGGGCTCAGGGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCA-- 212
                 ||||||  |||      |||  |||                  ||||||
Db        784 uLysGlySerProGlyAlaGlnGlyProPro-----------------GlyAlaProGl 798

Qy        211 -----------------------------GAAGCTGCGGCCTCTCCTCCTTGCTGCCG 183
                                           |||   ||||||
Db        798 yProLeuGlyIleAlaGlyIleThrGlyAlaArgGlyLeuAlaGlyProProGlyMetPr 818

Qy        182 CCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGC 123
                   ||||||      ||||||             :::              |||
Db        818 oGly---ProArgGlySerProGlyProGlnGlyValLysGlyGluSerGlyLysProGl 837

Qy        122 TTCTCAG-----------------------------CCCATGCTCAACACCTGCTGCTG 93
               :::                                |||          |||
Db        837 yAlaAsnGlyLeuSerGlyGluArgGlyProProGlyProGln-GlyLeuProGlyLeuA 857

Qy         92 TGGGGCACCTCAGTGGGGACACGTCTCATCACTCAGATCCTGGC--------------- 49
               |||      :::|||:::               :::||||||
Db        857 laGly---ThrAlaGlyGluProGlyArgAspGlyAsnProGlySerAspGlyLeuProG 876

Qy         48 ------------------------------CGAGGCGCGCGGCTGTCACCCGGAGCC 22
                                           ||||||         |||||||||||
Db        876 lyArgAspGlySerProGlyGlyLysGlyAspArgGlyGluAsnGlySerProGlyAla 895
```

RESULT 27
US-09-570-573-21
; Sequence 21, Application US/09570573
; Patent No. 6342361
; GENERAL INFORMATION:
; APPLICANT: Qvist, Per
; APPLICANT: Bonde, Martin
; TITLE OF INVENTION: A Method for Assaying Collagen Fragments
; TITLE OF INVENTION: in Body Fluids, A Test Kit and Means for Carrying Out the
; TITLE OF INVENTION: Method and Use of the Method to Diagnose the Presence of
; TITLE OF INVENTION: Disorders Associated with the Metabolism of
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/570,573
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/187,319
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Gogoris, Adda C
; REGISTRATION NUMBER: 29,714
; REFERENCE/DOCKET NUMBER: 4305/08701
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-527-7700
; TELEFAX: 212-753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1078 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: COLLAGEN ALPHA 1 (III)
US-09-570-573-21

Alignment Scores:

Pred. No.:	6.94e-20	Length:	1078
Score:	364.50	Matches:	290
Percent Similarity:	30.79%	Conservative:	57
Best Local Similarity:	25.73%	Mismatches:	354
Query Match:	5.86%	Indels:	426
DB:	4	Gaps:	62

US-09-759-143-110 (1-3410) x US-09-570-573-21 (1-1078)

```
Qy       2939 GGAAGTGGGGGGAACCAGGCTGGGCCAAGAGAAGAGGGGTGGTTAGGGAAGCCGTTGAGA 2880
              ||||||  |||  |||        |||  |||  |||         |||  |||
Db         48 GlySerProGlyTyrGlnGlyProProGlyGluProGlyGlnAlaGly---ProSerGly 66

Qy       2879 CCTGAAGCCCCCACCCTCTACCTTCCTTCAACACCCTAACCTTGGGTAACAGCATTTGGAA 2820
              |||    ||||||
Db         67 ProProGlyProProGly------------------------------------------ 72

Qy       2819 TTATCATTTGGGATGAGTAGAATTTCCAAGGTCCTGGGTTAGGCATTTTGGGGGGCCAGA 2760
                 :::  |||  |||      :::          |||   :::              |||
Db         73 ---AlaIleGlyProSerGlyProAlaGlyLysAspGlyGluSerGlyArgProGlyArg 91

Qy       2759 CCCCAGGAGAAGAAGATTCTGGCAATGATCAGCCCAATGACCAGCTATCTCAGGGGACCT 2700
              |||   ::::::   :::                |||          ::::::||||||
Db         92 ProGlyAspArgGlyLeu--------------ProGlyProProGlyIleLysGlyPro 106

Qy       2699 GATTGTTGGGGATCCCCCACCCTACCCAAATATTAGACACCAACACAGA----------- 2651
                       |||    |||   :::|||   ::::::     ||||||
Db        107 Ala------GlyIle-ProGlyPheProGlyMetLys---GlyHisArgGlyPheAspGl 123

Qy       2650 ------------AAAGCTAGCAATGGATTCCCTTCTACTTTGTTAAATAAATAAGTTAA 2604
                          |||        |||   |||                      :::||
Db        123 yArgAsnGlyGluLysGlyGluThrGlyAlaProGly-----------------LeuLy 137

Qy       2603 ATATTTAAATGCCTGTGTCTCTGTGATGGCAACAGAAGGACCAACAGGCCACATCCTGAT 2544
              |      |||                            ||||||    |||
Db        137 sGlyGluAsnGlyLeuProGlyGluAsnGlyAlaProGlyProMetGlyProArgGlyAl 157

Qy       2543 AAAAGGTAAGAGGGGGGTGGATCAGCAAAAAGACAGTGCTGTGGGCTGAGGGGACCTGGT 2484
                 |||:::||||||                    :::|||        |||:::
Db        157 aProGlyGluArgGlyArgProGlyLeuProGlyAlaAlaGlyAlaArgGlyAsn----- 175

Qy       2483 TCTTGTGTGTTGCCCCTCAGGACTCTTCCCCTACAAATAAGTCATATGTTCAAATCCCAT 2424

Db        175 ------------------------------------------------------------ 175

Qy       2423 GGAGGAGTGTTTCATCCTAGAAACTCCCATGCAAGAGCTACATTAAACGAAGCTGCAGGT 2364

Db        175 ------------------------------------------------------------ 175

Qy       2363 TAAGGGGCTTAGA------GATGGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAA 2310
                 |||   |||      ||||||:::||||||                  |||
Db        176 -AspGlyAlaArgGlySerAspGlyGlnProGly-----------------ProProGl 189

Qy       2309 CCCTTCTCTAGGTGTGTCTCAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCC 2250
               |||      |||   :::                              |||:::||
Db        189 yProPro---GlyThrAlaGlyPhePro----------------------GlySerPr 200
```

```
Qy   2363 TAAGGGGCTTAGA------GATGGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAA 2310
Db    176 -AspGlyAlaArgGlySerAspGlyGlnProGly-----------------ProProGl 189
Qy   2309 CCCTTCTCTAGGTGTGTCTCAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCC 2250
Db    189 yProPro---GlyThrAlaGlyPhePro-----------------------GlySerPr 200
Qy   2249 ACCTGCAGAGTCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCA 2190
Db    200 oGlyAlaLysGlyGluValGlyProAlaGlySerProGly------------------- 213
Qy   2189 GACTGAAACCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAG 2130
Db    214 ------------------------SerAsnGlyAlaPro------------GlyGlnAr 221
Qy   2129 AGGGACGCCCCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAG 2070
Db    221 gGlyGluProGlyProGlnGly------------HisAlaGlyAlaGlnGlyProProGl 237
Qy   2069 AGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGC 2010
Db    237 yProProGlyIleAsnGlySer--------------------------ProGlyGly-- 247
Qy   2009 TAACAGGAGCGGGGAGCTGGGACCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAG 1950
Db    248 -------LysGlyGluMetGlyProAla--GlyIleProGlyAlaPro----------- 260
Qy   1949 TTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGT 1890
Db    261 -----------GlyLeuMetGlyAlaArgGly-----------ProProGlyProAla 272
Qy   1889 AAATGGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGA 1830
Db    273 GlyAlaAsnGlyAlaProGlyLeuArg-------------------------GlyGly 283
Qy   1829 GCTGGACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCAC--- 1773
Db    284 AlaGlyGluProGlyLysAsnGlyAlaLysGlyGluProGlyProArgGlyGluArgGly 303
Qy   1772 -----------------TATCCAGGATGGCGAGGT--------------------CCA 1752
Db    304 GluAlaGlyIleProGlyValProGlyAlaLysGlyGluAspGlyLysAspGlySerPro 323
Qy   1751 GGCAGATGCCCCGGCCCGGAACCACCCTGGCCTCGGTGGGCTCACCCACCACCACACGTA 1692
Db    324 Gly-AspProGlyAlaAsnGlyLeuProGlyAlaAlaGly--------------GluAr 338
Qy   1691 CGGAGACATCACAGGCAGAGGCCCCGCA--------------------GAGCGCGGGTG 1653
Db    338 gGlyAlaLeuGlySerArgGlyProAlaGlyProAsnGlyIleProGlyGluLys-GlyP 358
Qy   1652 GAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGGG---AAGGGAGCTC 1596
Db    358 roAlaGlyGluArgGly---AlaProGlyProAlaGlyProArgGlyAlaAlaGlyGluP 377
Qy   1595 CAGGCTTAGGG------CCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCTAGCAC 1542
Db    377 roGlyArgAspGlyValProGly-GlyProGlyMetArgGlyMetPro----------- 392
Qy   1541 CTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGTAGAGGGAGG 1482
Db    393 --------------------------GlySerProGlyGlyProGlySerAspGlyLys 403
Qy   1481 CCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAG 1422
Db    404 Pro---GlyProProGlySer-GlnGlyGluSerGlyArgProGlyProProGlyProSe 422
Qy   1421 CTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACCGGCAGCCACAGGGAAAGCTG 1362
Db    422 r--GlyProArgGlyGlnProGlyVal---MetGlyPheProGlyProLysGlyAsn--- 439
Qy   1361 CCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCTGCACCAGCCGGTCCATGACCA 1302
Db    440 AspGlyAlaProGlyLysAsnGlyGlu------ArgGlyGlyProGlyGlyProGlyPro 457
Qy   1301 GAGAGAAGACCAGGGAGATGGCGCACTGCAGGA--------------------ACAGCC 1263
Db    458 GlnGlyProProGlyLysAsnGlyGluTyrGlyProGlnGlyProProGlyProThrGly 477
Qy   1262 CCAGGCTGCCCATCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAG 1203
Db    478 ProGlyGlyAspLys---------------GlyAspThrGlyProArgGlyProGlnGly 492
Qy   1202 CTCTGGGCACGCCCTGGTACA-------------------------------------- 1182
Db    493 LeuGlnGlyLeuProGlyThrGlyGlyProProGlyGluAsnGluLysProGlyGluPro 512
Qy   1181 ---------------------------------------------------------GCC 1179
Db    513 GlyProLysGlyGluAlaGlyAlaProGlyAlaProGlyGlyLysGlyAspAlaGlyAla 532
Qy   1178 CCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGCTGCACA 1119
Db    533 ProGlyGluArgGlyPro-------------------------ProGlyLeuAlaGly 543
Qy   1118 GCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACAGCTGGTGCAGCC 1059
Db    544 AlaProGlyLeuArgGlyGlyAlaGlyProGlyProGluGlyGlyLysGlyAlaAla 563
Qy   1058 GGGGAAGCAGGGCGCCCAGGTTCC------------------------------------ 1035
Db    564 GlyProProGlyProProGlyAlaAlaGlyThrProGlyLeuGlnGlyMetProGlyGlu 583
Qy   1034 -----------------------------------GGAAAGCCAAGCGGGCCCGGC 1014
Db    584 ArgGlyGlyLeuGlySerProGlyProLysGlyAspLysGlyGluProGlyGlyProGly 603
Qy   1013 ATGGACAGCAGTGGGGCGACA------AGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGG 960
Db    604 AlaAspGlyValProGlyLysAspGlyProArgGlyProThr----------------- 617
Qy    959 TGGGGCCCAGCGCTGCCTCCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGA 900
Db    618 ---GlyProIleGlyProProGlyProAla----------------------GlyGln 628
Qy    899 TGAGGGTGAGCAGGCCAAAGAGGCACTCCTCCTGGGTGCCCAGGT---AGGGGGCCAGGG 843
Db    629 ProGlyAspLysGlyGluGlyGlyAla---ProGlyLeuProGlyIleAlaGlyPro--- 646
Qy    842 CACTGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGACTGATCA 783
Db    647 -------------------------ArgGlySerProGlyGlu-------------- 652
Qy    782 TGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGT---CCCGGAAGAGGT 729
Db    653 ---ArgGlyGluThrGlyProProGlyProAlaGlyPheProGlyAlaProGlyGlnAsn 671
Qy    728 CAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCA 669
Db    672 GlyGluProGlyGlyLysGlyGluArgGlyAlaProGlyGluLysGlyGluGlyGlyPro 691
Qy    668 CGCCCAGGATGAGCAGTGCCAGCTCCAGGG---------GCCTGGGATCCGGGCACAGCA 618
Db    692 ---ProGlyValAlaValProProGlyGlySerGlyProAlaGlyProProGlyProGln 710
Qy    617 GCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCAAGGACA 558
Db    711 GlyValLysGly-----------------GluArgGlySerProGlyGlyPro------ 722
Qy    557 GTGCCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAGC 498
Db    723 ----------------------------------GlyAlaAlaGlyPhe-------- 727
Qy    497 CTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCATGGTCATGA 438
```

```
Qy        557 GTGCCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAGC 498
                                                  |||:::|||
Db        723 -----------------------------------GlyAlaAlaGlyPhe--------- 727

Qy        497 CTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCATGGTCATGA 438
                ||||||      |||||||||                        |||   |||
Db        728 ProGlyAlaArgGlyLeuProGlyPro--------------------ProGlySerAsn 740

Qy        437 ACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCCA 378
                :::|||  ||||||  |||:::
Db        741 GlyAsnProGlyProProGlyProSerGlySer-------------------------- 751

Qy        377 AACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGTGCC 318
                       ||||||:::       |||  |||     ||||||  :::||||||
Db        752 ---------ProGlyLys------AspGlyProProGlyProAlaGlyAsnThrGlyAla 766

Qy        317 GCAGCAGGCGGCTCACCCACAGCCTCTGGACCATAGTGGG-----------CCAGGCGG 270
                      |||||||||    |||||||||  |||         ||||||
Db        767 Pro------GlySerProGlyValSerGlyPro-LysGlyAspAlaGlyGlnProGlyGl 784

Qy        269 GTAGGGCTCAGGGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCA-- 212
                ||||||  |||      |||  |||                 ||||||
Db        784 uLysGlySerProGlyAlaGlnGlyProPro-----------------GlyAlaProGl 798

Qy        211 ------------------------------GAAGCTGCGGCCTCTCCTCCTTGCTGCCG 183
                                                |||  ||||||
Db        798 yProLeuGlyIleAlaGlyIleThrGlyAlaArgGlyLeuAlaGlyProProGlyMetPr 818

Qy        182 CCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGC 123
                      ||||||     ||||||            :::            |||
Db        818 oGly---ProArgGlySerProGlyProGlnGlyValLysGlyGluSerGlyLysProGl 837

Qy        122 TTCTCAG----------------------------CCCATGCTCAACACCTGCTGCTG 93
               :::                                 |||       |||
Db        837 yAlaAsnGlyLeuSerGlyGluArgGlyProProGlyProGln-GlyLeuProGlyLeuA 857

Qy         92 TGGGGCACCTCAGTGGGGACACGTCTCATCACTCAGATCCTGGC---------------- 49
                |||       :::|||:::               :::||||||
Db        857 laGly---ThrAlaGlyGluProGlyArgAspGlyAsnProGlySerAspGlyLeuProG 876

Qy         48 ------------------------------CGAGGCGCGCGGCTGTCACCCGGAGCC 22
                                             ||||||          ||||||||||||
Db        876 lyArgAspGlySerProGlyGlyLysGlyAspArgGlyGluAsnGlySerProGlyAla 895
```

RESULT 26
US-09-500-811-21
; Sequence 21, Application US/09500811
; Patent No. 6323314
; GENERAL INFORMATION:
; APPLICANT: Qvist, Per
; APPLICANT: Bonde, Martin
; TITLE OF INVENTION: A Method for Assaying Collagen Fragments
; TITLE OF INVENTION: in Body Fluids, A Test Kit and Means for Carrying Out the
; TITLE OF INVENTION: Method and Use of the Method to Diagnose the Presence of
; TITLE OF INVENTION: Disorders Associated with the Metabolism of
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/500,811
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/187,319
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Gogoris, Adda C
; REGISTRATION NUMBER: 29,714
; REFERENCE/DOCKET NUMBER: 4305/08701
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-527-7700
; TELEFAX: 212-753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1078 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: COLLAGEN ALPHA 1 (III)
US-09-500-811-21

Alignment Scores:

Pred. No.:	6.94e-20	Length:	1078
Score:	364.50	Matches:	290
Percent Similarity:	30.79%	Conservative:	57
Best Local Similarity:	25.73%	Mismatches:	354
Query Match:	5.86%	Indels:	426
DB:	4	Gaps:	62

US-09-759-143-110 (1-3410) x US-09-500-811-21 (1-1078)

```
Qy       2939 GGAAGTGGGGGAACCAGGCTGGGCCAAGAGAAGAGGGGTGGTTAGGGAAGCCGTTGAGA 2880
              ||||||   |||   |||          |||   |||   |||         |||   |||
Db         48 GlySerProGlyTyrGlnGlyProProGlyGluProGlyGlnAlaGly---ProSerGly 66

Qy       2879 CCTGAAGCCCCACCCTCTACCTTCCTTCAACACCCTAACCTTGGGTAACAGCATTTGGAA 2820
              |||      ||||||
Db         67 ProProGlyProProGly------------------------------------------ 72

Qy       2819 TTATCATTTGGGATGAGTAGAATTTCCAAGGTCCTGGGTTAGGCATTTTGGGGGGCCAGA 2760
                 :::   |||   |||         :::         |||   :::         |||
Db         73 ---AlaIleGlyProSerGlyProAlaGlyLysAspGlyGluSerGlyArgProGlyArg 91

Qy       2759 CCCCAGGAGAAGAAGATTCTGGCAATGATCAGCCCAATGACCAGCTATCTCAGGGGACCT 2700
              |||   ::::::   :::                   |||      ::::::||||||
Db         92 ProGlyAspArgGlyLeu---------------ProGlyProProGlyIleLysGlyPro 106

Qy       2699 GATTGTTGGGGATCCCCCACCCTACCCAAATATTAGACACCAACACAGA----------- 2651
                      |||   |||   :::|||   ::::::      ||||||
Db        107 Ala------GlyIle-ProGlyPheProGlyMetLys---GlyHisArgGlyPheAspGl 123

Qy       2650 ------------AAAGCTAGCAATGGATTCCCTTCTACTTTGTTAAATAAATAAGTTAA 2604
                           |||           |||   |||                   :::||
Db        123 yArgAsnGlyGluLysGlyGluThrGlyAlaProGly-----------------LeuLy 137

Qy       2603 ATATTTAAATGCCTGTGTCTCTGTGATGGCAACAGAAGGACCAACAGGCCACATCCTGAT 2544
              |    |||                                  ||||||   |||
Db        137 sGlyGluAsnGlyLeuProGlyGluAsnGlyAlaProGlyProMetGlyProArgGlyAl 157

Qy       2543 AAAAGGTAAGAGGGGGGTGGATCAGCAAAAAGACAGTGCTGTGGGCTGAGGGGACCTGGT 2484
                 |||:::||||||                  :::|||      |||:::
Db        157 aProGlyGluArgGlyArgProGlyLeuProGlyAlaAlaGlyAlaArgGlyAsn----- 175

Qy       2483 TCTTGTGTGTTGCCCCTCAGGACTCTTCCCCTACAAATAAGTCATATGTTCAAATCCCAT 2424
Db        175 ------------------------------------------------------------ 175

Qy       2423 GGAGGAGTGTTTCATCCTAGAAACTCCCATGCAAGAGCTACATTAAACGAAGCTGCAGGT 2364
Db        175 ------------------------------------------------------------ 175
```

```
Db      175 ------------------------------------------------------------ 175

Qy     2423 GGAGGAGTGTTTCATCCTAGAAACTCCCATGCAAGAGCTACATTAAACGAAGCTGCAGGT 2364

Db      175 ------------------------------------------------------------ 175

Qy     2363 TAAGGGGCTTAGA------GATGGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAA 2310
                |||   |||      ||||||:::||||||                   |||
Db      176 -AspGlyAlaArgGlySerAspGlyGlnProGly------------------ProProGl 189

Qy     2309 CCCTTCTCTAGGTGTGTCTCAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCC 2250
             |||    |||   :::                                 .|||:::||
Db      189 yProPro---GlyThrAlaGlyPhePro------------------------GlySerPr 200

Qy     2249 ACCTGCAGAGTCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCA 2190
             |   |||:::                  |||   ||||||
Db      200 oGlyAlaLysGlyGluValGlyProAlaGlySerProGly-------------------- 213

Qy     2189 GACTGAAACCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAG 2130
                                         :::|||:::|||                :::||
Db      214 ------------------------SerAsnGlyAlaPro------------GlyGlnAr 221

Qy     2129 AGGGACGCCCCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAG 2070
             ||||   |||   |||                        :::   |||||||||
Db      221 gGlyGluProGlyProGlnGly------------HisAlaGlyAlaGlnGlyProProGl 237

Qy     2069 AGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGC 2010
                |||         ||||||                           |||   |||
Db      237 yProProGlyIleAsnGlySer--------------------------ProGlyGly-- 247

Qy     2009 TAACAGGAGCGGGGAGCTGGGACCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAG 1950
                       ||||||:::||||||:::   |||    |||      |||
Db      248 -------LysGlyGluMetGlyProAla--GlyIleProGlyAlaPro------------ 260

Qy     1949 TTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGT 1890
                            :::      :::|||                |||        :::
Db      261 -----------GlyLeuMetGlyAlaArgGly------------ProProGlyProAla 272

Qy     1889 AAATGGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGA 1830
                              ||||||||||||                         |||
Db      273 GlyAlaAsnGlyAlaProGlyLeuArg-------------------------GlyGly 283

Qy     1829 GCTGGACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCAC--- 1773
             ||||||:::                      |||   ||||||    |||
Db      284 AlaGlyGluProGlyLysAsnGlyAlaLysGlyGluProGlyProArgGlyGluArgGly 303

Qy     1772 ------------------TATCCAGGATGGCGAGGT--------------------CCA 1752
                                ||||||   :::|||                     |||
Db      304 GluAlaGlyIleProGlyValProGlyAlaLysGlyGluAspGlyLysAspGlySerPro 323

Qy     1751 GGCAGATGCCCCGGCCCGGAACCACCCTGGCCTCGGTGGGCTCACCCACCACCACACGTA 1692
             ||| |||     |||         ||||||     |||
Db      324 Gly-AspProGlyAlaAsnGlyLeuProGlyAlaAlaGly--------------GluAr 338

Qy     1691 CGGAGACATCACAGGCAGAGGCCCCGCA--------------------GAGCGCGGGTG 1653
              |||    :::       |||||||||||                   |||::: |||
Db      338 gGlyAlaLeuGlySerArgGlyProAlaGlyProAsnGlyIleProGlyGluLys-GlyP 358

Qy     1652 GAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGGG---AAGGGAGCTC 1596
                |||   |||           |||   |||    |·|   |||       |||    |
Db      358 roAlaGlyGluArgGly---AlaProGlyProAlaGlyProArgGlyAlaAlaGlyGluP 377

Qy     1595 CAGGCTTAGGG------CCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCTAGCAC 1542
             |||||            |||||| |||                |||    |||
Db      377 roGlyArgAspGlyValProGly-GlyProGlyMetArgGlyMetPro----------- 392

Qy     1541 CTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGTAGAGGGAGG 1482
                                     |||:::|||       ||||||       |||:::
Db      393 -------------------------GlySerProGlyGlyProGlySerAspGlyLys 403

Qy     1481 CCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAG 1422
             |||              |||||| ||||||        |||    ||||||       |||   ||

Db      404 Pro---GlyProProGlySer-GlnGlyGluSerGlyArgProGlyProProGlyProSe 422

Qy     1421 CTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACCGGCAGCCACAGGGAAAGCTG 1362
             |    ||||||        |||     |||              |||:::|||
Db      422 r--GlyProArgGlyGlnProGlyVal---MetGlyPheProGlyProLysGlyAsn--- 439

Qy     1361 CCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCTGCACCAGCCGGTCCATGACCA 1302
                      |||   :::      |||                 |||   ||||||   |||
Db      440 AspGlyAlaProGlyLysAsnGlyGlu------ArgGlyGlyProGlyGlyProGlyPro 457

Qy     1301 GAGAGAAGACCAGGGAGATGGCGCACTGCAGGA--------------------ACAGCC 1263
             :::      ||||||:::            |||                       |||
Db      458 GlnGlyProProGlyLysAsnGlyGluTyrGlyProGlnGlyProProGlyProThrGly 477

Qy     1262 CCAGGCTGCCCATCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAG 1203
             ||||||                               :::|||||||||   |||
Db      478 ProGlyGlyAspLys---------------GlyAspThrGlyProArgGlyProGlnGly 492

Qy     1202 CTCTGGGCACGCCCTGGTACA------------------------------------- 1182
             |||       . |||||||||
Db      493 LeuGlnGlyLeuProGlyThrGlyGlyProProGlyGluAsnGluLysProGlyGluPro 512

Qy     1181 -------------------------------------------------------GCC 1179
                                                                    |||
Db      513 GlyProLysGlyGluAlaGlyAlaProGlyAlaProGlyGlyLysGlyAspAlaGlyAla 532

Qy     1178 CCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGCTGCACA 1119
             |||      |||   |||                          |||
Db      533 ProGlyGluArgGlyPro--------------------------ProGlyLeuAlaGly 543

Qy     1118 GCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACAGCTGGTGCAGCC 1059
             |||       |||   |||||||||                     |||||||||
Db      544 AlaProGlyLeuArgGlyGlyAlaGlyProProGlyProGluGlyGlyLysGlyAlaAla 563

Qy     1058 GGGGAAGCAGGGCGCCCAGGTTCC------------------------------------ 1035
             |||       |||   ||||||:::
Db      564 GlyProProGlyProProGlyAlaAlaGlyThrProGlyLeuGlnGlyMetProGlyGlu 583

Qy     1034 --------------------------------------GGAAAGCCAAGCGGCCCGGC 1014
                                                   |||:::|||   |||||||||
Db      584 ArgGlyGlyLeuGlySerProGlyProLysGlyAspLysGlyGluProGlyGlyProGly 603

Qy     1013 ATGGACAGCAGTGGGGCGACA------AGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGG 960
             . |||                     ||||||||||||
Db      604 AlaAspGlyValProGlyLysAspGlyProArgGlyProThr------------------ 617

Qy      959 TGGGGCCCAGCGCTGCCTCCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGA 900
                ||||||    ||||||   |||                        . |||:::
Db      618 ---GlyProIleGlyProProGlyProAla------------------------GlyGln 628

Qy      899 TGAGGGTGAGCAGGCCAAAGAGGCACTCCTCCTGGGTGCCCAGGT---AGGGGGCCAGGG 843
               |||    |||:::   |||      ||||||   ||||||       ||||||
Db      629 ProGlyAspLysGlyGluGlyGlyAla---ProGlyLeuProGlyIleAlaGlyPro--- 646

Qy      842 CACTGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGACTGATCA 783
                                         ||||||||||||
Db      647 --------------------------ArgGlySerProGlyGlu--------------- 652

Qy      782 TGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGT---CCCGGAAGAGGT 729
               |||                |||||| :::|||   ||||||   ||||||:::
Db      653 ---ArgGlyGluThrGlyProProGlyProAlaGlyPheProGlyAlaProGlyGlnAsn 671

Qy      728 CAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCA 669
                  |||        |||        ||||||   :::                |||
Db      672 GlyGluProGlyGlyLysGlyGluArgGlyAlaProGlyGluLysGlyGluGlyGlyPro 691

Qy      668 CGCCCAGGATGAGCAGTGCCAGCTCCAGGG---------GCCTGGGATCCGGGCACAGCA 618
               ||||||   |||||||||      |||          |||      ||||||
Db      692 ---ProGlyValAlaValProProGlyGlySerGlyProAlaGlyProProGlyProGln 710

Qy      617 GCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCAAGGACA 558
               :::                           |||||||||||   |||   |||
Db      711 GlyValLysGly------------------GluArgGlySerProGlyGlyPro------ 722
```

```
Qy       1409 GCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTG 1468
               |||      |||      |||   |||||||||:::|||     :::|||
Db        481 yPro------GlnGlyLeuProGlySerProGlyAlaProGly-ThrProGlyPro---- 497

Qy       1469 CCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCT------------- 1515
                              |||||||||       ||||||      |||   |||
Db        498 --------GlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGluGlyGlnG 515

Qy       1516 -----------------GCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGAC 1558
                                |||  |||       ||||||   |||          :::
Db        515 lnHisHisLeuGlyGlyAla-LysGlnAlaGlyAspValGlySerProGlyAlaProGly 534

Qy       1559 AGCCTGATGACCAGCTTCCTGCCAGGCCCT------AAGCCTGGAGCTCCCTTCCCTAAT 1612
              :::                ||||||||              ||||||    |||   |||
Db        535 ThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGln 554

Qy       1613 GGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCC 1672
               |||      |||:::     |||:::          |||    |||       |||   |||:::
Db        555 GlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySer 574

Qy       1673 ---------------------------------------------------------TCT 1675

Db        575 ProGlyAlaProGlyThrProGlyGluGlyGlnGlnHisHisLeuGlyGlyAlaLysGln 594

Qy       1676 GCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCG 1735
               |||   |||                    |||||| ||||||      ||| |||      |
Db        595 AlaGlyAsp-----------------ValGly-SerProGlyAlaPro-GlyThrProG 608

Qy       1736 GGCCGGGGCATCTGCCTGGA------------CCTCGCCATCCTGGA---TAGTGCCTTC 1780
               |||||     |||||||||             |||      ||||||          ||||
Db        608 lyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuP 628

Qy       1781 CTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCT 1840
               ||      ||||||   |||                          |||      |||
Db        628 roGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro---- 646

Qy       1841 GTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGC------CATTTACTTT 1894
                            |||             ||||||   |||:::         ||||||
Db        647 -------------GlyAlaProGlyThrProGlyGluGlyGlnGlnHisHisLeu---- 660

Qy       1895 GCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCC 1954
                      |||                  ::::::   |||     :::    |||
Db        661 --------GlyGly------AlaLysGlnAlaGlyAspValGlySerProGlyAlaProG 676

Qy       1955 AGCACATTGGGGTGGAGGGCCTGCCTCACTGGTCCCAGCTCCCCGCTCCTGTTAGCCCC 2014
                         |||   :::|||||||||          |||  |||             |||
Db        676 lyThrProGlyProGlnGlyLeuProGlySerProGlyAla-ProGlyThrProGlyPro 695

Qy       2015 ATGGGGCTGCCGGGCTGGCCG 2035
                 ||||||||||||    |||
Db        696 GlnGlyLeuProGlySerPro 702
```

RESULT 25
US-08-963-825-21
; Sequence 21, Application US/08963825
; Patent No. 6110689
; GENERAL INFORMATION:
; APPLICANT: Qvist, Per
; APPLICANT: Bonde, Martin
; TITLE OF INVENTION: A Method for Assaying Collagen Fragments
; TITLE OF INVENTION: in Body Fluids, A Test Kit and Means for Carrying Out the
; TITLE OF INVENTION: Method and Use of the Method to Diagnose the Presence of
; TITLE OF INVENTION: Disorders Associated with the Metabolism of
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/963,825
; FILING DATE:
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/187,319
; FILING DATE: 21-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Gogoris, Adda C
; REGISTRATION NUMBER: 29,714
; REFERENCE/DOCKET NUMBER: 4305/08701
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-527-7700
; TELEFAX: 212-753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1078 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: COLLAGEN ALPHA 1 (III)
US-08-963-825-21

```
Alignment Scores:
Pred. No.:                 6.94e-20      Length:         1078
Score:                     364.50        Matches:        290
Percent Similarity:        30.79%        Conservative:   57
Best Local Similarity:     25.73%        Mismatches:     354
Query Match:               5.86%         Indels:         426
DB:                        3             Gaps:           62
```

US-09-759-143-110 (1-3410) x US-08-963-825-21 (1-1078)

```
Qy       2939 GGAAGTGGGGGGAACCAGGCTGGGCCAAGAGAAGAGGGGTGGTTAGGGAAGCCGTTGAGA 2880
                ||||||   |||    |||      |||   |||   |||       |||   |||
Db         48 GlySerProGlyTyrGlnGlyProProGlyGluProGlyGlnAlaGly---ProSerGly 66

Qy       2879 CCTGAAGCCCCACCCTCTACCTTCCTTCAACACCCTAACCTTGGGTAACAGCATTTGGAA 2820
               |||      ||||||
Db         67 ProProGlyProProGly------------------------------------------ 72

Qy       2819 TTATCATTTGGGATGAGTAGAATTTCCAAGGTCCTGGGTTAGGCATTTTGGGGGGCCAGA 2760
                 :::   |||   |||      :::          |||    :::            |||
Db         73 ---AlaIleGlyProSerGlyProAlaGlyLysAspGlyGluSerGlyArgProGlyArg 91

Qy       2759 CCCCAGGAGAAGAAGATTCTGGCAATGATCAGCCCAATGACCAGCTATCTCAGGGGACCT 2700
               |||   ::::::   :::                |||              :::::||||||
Db         92 ProGlyAspArgGlyLeu---------------ProGlyProProGlyIleLysGlyPro 106

Qy       2699 GATTGTTGGGGATCCCCCACCCTACCCAAATATTAGACACCAACACAGA----------- 2651
                       |||     |||    :::|||    ::::::       ||||||
Db        107 Ala------GlyIle-ProGlyPheProGlyMetLys---GlyHisArgGlyPheAspGl 123

Qy       2650 -------------AAAGCTAGCAATGGATTCCCTTCTACTTTGTTAAATAAATAAGTTAA 2604
                             |||         |||    |||                   :::||
Db        123 yArgAsnGlyGluLysGlyGluThrGlyAlaProGly------------------LeuLy 137

Qy       2603 ATATTTAAATGCCTGTGTCTCTGTGATGGCAACAGAAGGACCAACAGGCCACATCCTGAT 2544
                |     |||                                   ||||||   |||
Db        137 sGlyGluAsnGlyLeuProGlyGluAsnGlyAlaProGlyProMetGlyProArgGlyAl 157

Qy       2543 AAAAGGTAAGAGGGGGGTGGATCAGCAAAAAGACAGTGCTGTGGGCTGAGGGGACCTGGT 2484
                    |||:::||||||                     :::|||       |||:::
Db        157 aProGlyGluArgGlyArgProGlyLeuProGlyAlaAlaGlyAlaArgGlyAsn----- 175

Qy       2483 TCTTGTGTGTTGCCCCTCAGGACTCTTCCCCTACAAATAAGTCATATGTTCAAATCCCAT 2424
```

```
;    TITLE OF INVENTION:  PRODUCTS COMPRISING SUBSTRATESCAPABLE
;    TITLE OF INVENTION:  OF ENZYMATIC CROSS-LINKING
;    NUMBER OF SEQUENCES:  105
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/397,633A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rowland, Bertram I
;      REGISTRATION NUMBER:  20,015
;      REFERENCE/DOCKET NUMBER:  A-58848-1/BIR PROP-011-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 781-1989
;      TELEFAX:  (415) 398-3249
;      TELEX:  910 277299
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  762 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-397-633A-26

Alignment Scores:
Pred. No.:              5.62e-20       Length:          762
Score:                  365.00         Matches:         226
Percent Similarity:     35.02%         Conservative:    37
Best Local Similarity:  30.09%         Mismatches:      292
Query Match:            5.69%          Indels:          197
DB:                     1              Gaps:            46

US-09-759-143-110 (1-3410) x US-08-397-633A-26 (1-762)

Qy       84 GGTGCCCCACAGCAGCAGGTGTTG------AGCATGGGCTGAGAAGCTGGACCGGCACCA 137
Db       52 GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro 71

Qy      138 AAGGGC------TGGCAGAAATGGGCGCCTGGC--------------------TGATTC 170
Db       72 GlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGluGlyGlnGlnHisHis 91

Qy      171 CTAGGCAGTTGGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGA 230
Db       92 LeuGlyGly---AlaLysGlnAlaGlyAspValGlySerProGlyAlaProGlyThrPro 110

Qy      231 AGCAGTTCTGGAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGT-- 288
Db      111 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro-GlnGlyLe 130

Qy      289 ---------CCAGAGGCTGTGGGTGAGCCG---CCTGCTGCGGCACCGGAAAGCCCAGC 335
Db      130 uProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGl 150

Qy      336 TCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCT 395
Db      150 yAlaProGlyThrProGlyGlu--------------------GlyGlnGlnHisHisLe 163

Qy      396 ATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGG 455
Db      163 uGlyGlyAlaLysGlnAlaGlyAspValGly---------------------------- 173

Qy      456 GCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACT 515
Db      174 -------SerProGlyAlaProGly---ThrProGlyProGln---------------- 184

Qy      516 GGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGC 575
Db      185 -------------------------------------GlyLeuProGlySerProGl 191

Qy      576 TGAGCCT---CTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCT------GTGCC 626
Db      191 yAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPr 211

Qy      627 CGGATCCCAG-------------GCCCCTGGAGCT---GGCACTGCTCATCCTGGGCGTGG 671
Db      211 oGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr------ProGlyGluGl 229

Qy      672 G-----------GCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCC 719
Db      229 yGlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAlaPr 249

Qy      720 TGCTCTCTGACCTCTTCCGGGACCC----------------------GGACCACTGTC 755
Db      249 oGly---------ThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyTh 266

Qy      756 GCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGCTGCCTGGGCTACCTCC 815
Db      266 rProGly-------------------------ProGlnGlyLeuProGlySerProGl 277

Qy      816 TGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCC------CTACCTGGGCACCCAGG 869
Db      277 yAla---------------ProGlyThrProGlyProGlnGlyLeuProGlySerPro-- 291

Qy      870 AGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGC 929
Db      292 -GlyAlaProGlyThrProGlyGluGlyGlnGlnHisHisLeuGlyGlyAlaLysGlnAl 311

Qy      930 TGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCT 989
Db      311 aGlyAspValGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySe 331

Qy      990 CCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCG---GAACCTGGGCG 1046
Db      331 rProGlyAlaPro------GlyThrProGlyProGlnGlyLeuProGlySerProGlyAl 349

Qy     1047 CCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCT 1106
Db      349 aProGlyThrPro---GlyProGlnGlyLeuProGlySerPro---GlyAlaProGlyTh 367

Qy     1107 TCGTGGCTGAGCTGTGCAG------CTGGATGGCACTCATGACCTTCACGCTGTT----- 1155
Db      367 rProGlyGluGlyGlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySe 387

Qy     1156 -------TTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGC 1208
Db      387 rProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla---Pr 406

Qy     1209 CGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCT-- 1266
Db      406 oGlyThrProGlyProGlnGlyLeu---------------ProGlySerProGlyAlaPr 421

Qy     1267 ----GTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGC 1322
Db      421 oGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGl 441

Qy     1323 AGCGATTCGGCAC---------TCGAGCAGTCTATTTGGCCAGTGTGGCAGCTT----TC 1369
Db      441 uGlyGlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAl 461

Qy     1370 CCTGTGGCTGCCGGTGCCACATGCCTGTCCC--------------------ACAGTGTG 1408
Db      461 aProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGl 481
```

```
Qy    231 AGCAGTTCTGGAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGT-- 288
Db    111 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro-GlnGlyLe 130
Qy    289 ----------CCAGAGGCTGTGGGTGAGCCG---CCTGCTGCGGCACCGGAAAGCCCAGC 335
Db    130 uProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGl 150
Qy    336 TCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCT 395
Db    150 yAlaProGlyThrProGlyGlu--------------------GlyGlnGlnHisHisLe 163
Qy    396 ATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGG 455
Db    163 uGlyGlyAlaLysGlnAlaGlyAspValGly----------------------------- 173
Qy    456 GCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACT 515
Db    174 -------SerProGlyAlaProGly---ThrProGlyProGln----------------- 184
Qy    516 GGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGC 575
Db    185 ----------------------------------------GlyLeuProGlySerProGl 191
Qy    576 TGAGCCT---CTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCT------GTGCC 626
Db    191 yAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPr 211
Qy    627 CGGATCCCAG-----------GCCCCTGGAGCT---GGCACTGCTCATCCTGGGCGTGG 671
Db    211 oGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr------ProGlyGluGl 229
Qy    672 G-----------GCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCC 719
Db    229 yGlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAlaPr 249
Qy    720 TGCTCTCTGACCTCTTCCGGGACCC-----------------------GGACCACTGTC 755
Db    249 oGly---------ThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyTh 266
Qy    756 GCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCC 815
Db    266 rProGly-------------------------ProGlnGlyLeuProGlySerProGl 277
Qy    816 TGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCC------CTACCTGGGCACCCAGG 869
Db    277 yAla--------------ProGlyThrProGlyProGlnGlyLeuProGlySerPro-- 291
Qy    870 AGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGC 929
Db    292 -GlyAlaProGlyThrProGlyGluGlyGlnGlnHisHisLeuGlyGlyAlaLysGlnAl 311
Qy    930 TGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCT 989
Db    311 aGlyAspValGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySe 331
Qy    990 CCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCG---GAACCTGGGCG 1046
Db    331 rProGlyAlaPro------GlyThrProGlyProGlnGlyLeuProGlySerProGlyAl 349
Qy   1047 CCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCT 1106
Db    349 aProGlyThrPro---GlyProGlnGlyLeuProGlySerPro---GlyAlaProGlyTh 367
Qy   1107 TCGTGGCTGAGCTGTGCAG------CTGGATGGCACTCATGACCTTCACGCTGTT----- 1155
Db    367 rProGlyGluGlyGlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySe 387
Qy   1156 -------TTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGC 1208
Db    387 rProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla---Pr 406
Qy   1209 CGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCT-- 1266
Db    406 oGlyThrProGlyProGlnGlyLeu--------------ProGlySerProGlyAlaPr 421
Qy   1267 ----GTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGC 1322
Db    421 oGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGl 441
Qy   1323 AGCGATTCGGCAC---------TCGAGCAGTCTATTTGGCCAGTGTGGCAGCTT----TC 1369
Db    441 uGlyGlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAl 461
Qy   1370 CCTGTGGCTGCCGGTGCCACATGCCTGTCCC----------------------ACAGTGTG 1408
Db    461 aProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGl 481
Qy   1409 GCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTG 1468
Db    481 yPro------GlnGlyLeuProGlySerProGlyAlaProGly-ThrProGlyPro---- 497
Qy   1469 CCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCT------------- 1515
Db    498 --------GlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGluGlyGlnG 515
Qy   1516 -----------------GCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGAC 1558
Db    515 lnHisHisLeuGlyGlyAla-LysGlnAlaGlyAspValGlySerProGlyAlaProGly 534
Qy   1559 AGCCTGATGACCAGCTTCCTGCCAGGCCCT------AAGCCTGGAGCTCCCTTCCCTAAT 1612
Db    535 ThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGln 554
Qy   1613 GGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCC 1672
Db    555 GlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySer 574
Qy   1673 --------------------------------------------------------TCT 1675
Db    575 ProGlyAlaProGlyThrProGlyGluGlyGlnGlnHisHisLeuGlyGlyAlaLysGln 594
Qy   1676 GCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCG 1735
Db    595 AlaGlyAsp-----------------ValGly-SerProGlyAlaPro-GlyThrProG 608
Qy   1736 GGCCGGGGCATCTGCCTGGA------------CCTCGCCATCCTGGA---TAGTGCCTTC 1780
Db    608 lyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuP 628
Qy   1781 CTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCT 1840
Db    628 roGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro---- 646
Qy   1841 GTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGC------CATTTACTTT 1894
Db    647 --------------GlyAlaProGlyThrProGlyGluGlyGlnGlnHisHisLeu---- 660
Qy   1895 GCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCC 1954
Db    661 --------GlyGly------AlaLysGlnAlaGlyAspValGlySerProGlyAlaProG 676
Qy   1955 AGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCC 2014
Db    676 lyThrProGlyProGlnGlyLeuProGlySerProGlyAla-ProGlyThrProGlyPro 695
Qy   2015 ATGGGGCTGCCGGGCTGGCCG 2035
Db    696 GlnGlyLeuProGlySerPro 702
```

RESULT 24
US-08-397-633A-26
; Sequence 26, Application US/08397633A
; Patent No. 5773577
; GENERAL INFORMATION:
; APPLICANT: Cappello, Joseph

```
Qy       1047 CCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCT 1106
               |||   :::|||       |||     ||||||  :::|||      ||||||
Db        348 aProGlyThrPro---GlyProGlnGlyLeuProGlySerPro---GlyAlaProGlyTh 366

Qy       1107 TCGTGGCTGAGCTGTGCAG------CTGGATGGCACTCATGACCTTCACGCTGTT----- 1155
                 |||       |||     |||  |||                  |||
Db        366 rProGlyGluGlyGlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySe 386

Qy       1156 -------TTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGC 1208
                            |||     |||         :::||||||        :::
Db        386 rProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla---Pr 405

Qy       1209 CGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGCAGCCTGGGGCT-- 1266
               |||     ||||||:::   |||                     |||   |||||||||
Db        405 oGlyThrProGlyProGlnGlyLeu--------------ProGlySerProGlyAlaPr 420

Qy       1267 ----GTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGC 1322
                     |||      :::  |||||||||                |||      |||
Db        420 oGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGl 440

Qy       1323 AGCGATTCGGCAC---------TCGAGCAGTCTATTTGGCCAGTGTGGCAGCTT----TC 1369
                    :::|||            :::        |||     ||||||       ::
Db        440 uGlyGlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAl 460

Qy       1370 CCTGTGGCTGCCGGTGCCACATGCCTGTCCC--------------------ACAGTGTG 1408
              :         |||    |||         |||                    |||
Db        460 aProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGl 480

Qy       1409 GCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTG 1468
               |||       |||       |||    |||||||||:::|||     :::|||
Db        480 yPro------GlnGlyLeuProGlySerProGlyAlaProGly-ThrProGlyPro---- 496

Qy       1469 CCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCT------------ 1515
                           |||||||||      ||||||      |||    |||
Db        497 --------GlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGluGlyGlnG 514

Qy       1516 ----------------GCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGAC 1558
                                 ||| |||       ||||||   |||          :::
Db        514 lnHisHisLeuGlyGlyAla-LysGlnAlaGlyAspValGlySerProGlyAlaProGly 533

Qy       1559 AGCCTGATGACCAGCTTCCTGCCAGGCCCT------AAGCCTGGAGCTCCCTTCCCTAAT 1612
              :::                  |||||||||            ||||||   |||   |||
Db        534 ThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGln 553

Qy       1613 GGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCC 1672
               |||     |||:::   |||:::         |||    |||       |||   |||:::
Db        554 GlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySer 573

Qy       1673 -----------------------------------------------------TCT 1675

Db        574 ProGlyAlaProGlyThrProGlyGluGlyGlnGlnHisHisLeuGlyGlyAlaLysGln 593

Qy       1676 GCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCG 1735
               |||    |||                   ||||||  ||||||      |||  |||      |
Db        594 AlaGlyAsp-----------------ValGly-SerProGlyAlaPro-GlyThrProG 607

Qy       1736 GGCCGGGGCATCTGCCTGGA-----------CCTCGCCATCCTGGA---TAGTGCCTTC 1780
               |||||    |||||||||               |||     ||||||          ||||
Db        607 lyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuP 627

Qy       1781 CTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCT 1840
               ||       ||||||   |||                              |||     |||
Db        627 roGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro---- 645

Qy       1841 GTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGC------CATTTACTTT 1894
                             |||           ||||||  |||:::            ||||||
Db        646 -------------GlyAlaProGlyThrProGlyGluGlyGlnGlnHisHisLeu---- 659

Qy       1895 GCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCC 1954
                          |||           ::::::  |||   :::   |||
Db        660 --------GlyGly------AlaLysGlnAlaGlyAspValGlySerProGlyAlaProG 675

Qy       1955 AGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCC 2014
                      |||    :::|||||||||                |||  |||          |||
Db        675 lyThrProGlyProGlnGlyLeuProGlySerProGlyAla-ProGlyThrProGlyPro 694

Qy       2015 ATGGGGCTGCCGGGCTGGCCG 2035
                    ||||||||||||    |||
Db        695 GlnGlyLeuProGlySerPro 701
```

RESULT 23
US-08-642-255-114
; Sequence 114, Application US/08642255
; Patent No. 5773249
; GENERAL INFORMATION:
;   APPLICANT: CAPPELLO, Joseph
;   APPLICANT: FERRARI, Franco A.
;   TITLE OF INVENTION: High Molecular Weight Collagen-Like
;   TITLE OF INVENTION: Protein Polymers
;   NUMBER OF SEQUENCES: 135
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
;   STREET: 4 Embarcadero Center, Suite 3400
;   CITY: San Francisco
;   STATE: California
;   COUNTRY: USA
;   ZIP: 94111-4187
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/642,255
;   FILING DATE:
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: ROWLAND, Bertram I.
;   REGISTRATION NUMBER: 20,015
;   REFERENCE/DOCKET NUMBER: A55556-3/BIR
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415) 494-8700
;   TELEFAX: (415) 494-8771
;   TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 114:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 762 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-642-255-114

Alignment Scores:

Pred. No.:	5.62e-20	Length:	762
Score:	365.00	Matches:	226
Percent Similarity:	35.02%	Conservative:	37
Best Local Similarity:	30.09%	Mismatches:	292
Query Match:	5.69%	Indels:	197
DB:	1	Gaps:	46

US-09-759-143-110 (1-3410) x US-08-642-255-114 (1-762)

```
Qy         84 GGTGCCCCACAGCAGCAGGTGTTG------AGCATGGGCTGAGAAGCTGGACCGGCACCA 137
               |||   |||        |||   |||         |||   |||               |||   |||
Db         52 GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro 71

Qy        138 AAGGGC------TGGCAGAAATGGGCGCCTGGC-------------------TGATTC 170
               :::|||                  |||||||||
Db         72 GlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGluGlyGlnGlnHisHis 91

Qy        171 CTAGGCAGTTGGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGA 230
               ||||||             :::|||   |||:::        :::    ||||||
Db         92 LeuGlyGly---AlaLysGlnAlaGlyAspValGlySerProGlyAlaProGlyThrPro 110
```

```
Db        662 erProGlyAlaMet-AspProGlyArg 670

RESULT 22
US-08-707-237A-84
; Sequence 84, Application US/08707237A
; Patent No. 5830713
;  GENERAL INFORMATION:
;    APPLICANT:  Ferrari, Franco A.
;    APPLICANT:  Capello, Joseph
;    APPLICANT:  Crissman, John W.
;    APPLICANT:  Dorman, Mary A.
;    TITLE OF INVENTION:  METHODS FOR PREPARING SYNTHETIC
;    TITLE OF INVENTION:  REPETITIVE DNA
;    NUMBER OF SEQUENCES:  108
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herbert
;      STREET:  Four Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/707,237A
;      FILING DATE:  03-SEP-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/175,155
;      FILING DATE:  29-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/053,049
;      FILING DATE:  22-APR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/609,716
;      FILING DATE:  06-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/269,429
;      FILING DATE:  09-NOV-1988
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/114,618
;      FILING DATE:  29-OCT-1987
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 06/927,258
;      FILING DATE:  04-NOV-1986
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Trecartin, Richard F.
;      REGISTRATION NUMBER:  31,801
;      REFERENCE/DOCKET NUMBER:  A-55186-10/WHD
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 781-1989
;      TELEFAX:  (415) 398-3249
;      TELEX:  910 277299
;  INFORMATION FOR SEQ ID NO:  84:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  761 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-707-237A-84

Alignment Scores:
Pred. No.:              5.61e-20      Length:         761
Score:                  365.00        Matches:        226
Percent Similarity:     35.02%        Conservative:   37
Best Local Similarity:  30.09%        Mismatches:     292
Query Match:            5.69%         Indels:         197
DB:                     2             Gaps:           46
```

```
US-09-759-143-110 (1-3410) x US-08-707-237A-84 (1-761)

Qy     84 GGTGCCCCACAGCAGCAGGTGTTG------AGCATGGGCTGAGAAGCTGGACCGGCACCA 137
          |||  |||    |||  |||     |||  |||       |||  |||
Db     51 GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro 70

Qy    138 AAGGGC------TGGCAGAAATGGGCGCCTGGC--------------------TGATTC 170
          :::|||                 |||||||||
Db     71 GlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGluGlyGlnGlnHisHis 90

Qy    171 CTAGGCAGTTGGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGA 230
          ||||||         :::|||  |||:::        :::   ||||||
Db     91 LeuGlyGly---AlaLysGlnAlaGlyAspValGlySerProGlyAlaProGlyThrPro 109

Qy    231 AGCAGTTCTGGAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGT-- 288
                      |||:::|||      |||      |||     |||||||||   |||
Db    110 GlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro-GlnGlyLe 129

Qy    289 ----------CCAGAGGCTGTGGGTGAGCCG---CCTGCTGCGGCACCGGAAAGCCCAGC 335
                    |||   |||   |||   |||   |||         |||   ||||||
Db    129 uProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGl 149

Qy    336 TCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCT 395
                |||    |||                                 |||::::::||||||||
Db    149 yAlaProGlyThrProGlyGlu--------------------GlyGlnGlnHisHisLe 162

Qy    396 ATGTGCCGCCTCTGCTGCTGGAAGTGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGG 455
          |     |||      ||||||      |||
Db    162 uGlyGlyAlaLysGlnAlaGlyAspValGly---------------------------- 172

Qy    456 GCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACT 515
                |||         ||||||     |||    |||:::
Db    173 -------SerProGlyAlaProGly---ThrProGlyProGln---------------- 183

Qy    516 GGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGC 575
                                                  :::   |||   |||
Db    184 --------------------------------------GlyLeuProGlySerProGl 190

Qy    576 TGAGCCT---CTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCT------GTGCC 626
                |||            |||:::|||        |||      |||      ||
Db    190 yAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPr 210

Qy    627 CGGATCCCAG-----------GCCCCTGGAGCT---GGCACTGCTCATCCTGGGCGTGG 671
          ||||  |||            :::|||||||||  ||||||      ||||||   ||
Db    210 oGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr------ProGlyGluGl 228

Qy    672 G-----------GCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCC 719
          |                 |||             |||   :::   |||   |||   ||
Db    228 yGlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAlaPr 248

Qy    720 TGCTCTCTGACCTCTTCCGGGACCC-----------------------GGACCACTGTC 755
          |               |||||||||                         |||   ::
Db    248 oGly---------ThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyTh 265

Qy    756 GCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCC 815
          :||||||                                  ||||||||||||   |||
Db    265 rProGly--------------------------ProGlnGlyLeuProGlySerProGl 276

Qy    816 TGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCC------CTACCTGGGCACCCAGG 869
            |||                    |||||||||         |||||||||   |||
Db    276 yAla---------------ProGlyThrProGlyProGlnGlyLeuProGlySerPro-- 290

Qy    870 AGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGC 929
            |||   |||      |||                   ||||||       :::     ||
Db    291 -GlyAlaProGlyThrProGlyGluGlyGlnGlnHisHisLeuGlyGlyAlaLysGlnAl 310

Qy    930 TGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCT 989
            ||||      |||             |||                :::   :::
Db    310 aGlyAspValGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySe 330

Qy    990 CCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCG---GAACCTGGGCG 1046
                ||||||              |||||||||   |||   |||      ||||||
Db    330 rProGlyAlaPro------GlyThrProGlyProGlnGlyLeuProGlySerProGlyAl 348
```

```
Qy   2045 ---------------CAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAACAGGAG 2001
Db    100 GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro 119
Qy   2000 CGGGGAGCTGGGACCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTTCTACG 1941
Db    120 GlnGlyLeuProGlySerProGlyAlaProGlyThrPro-------------------- 132
Qy   1940 CTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGA 1881
Db    133 --------------------GlyProGlnGlyLeuPro-------------------- 138
Qy   1880 CCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTGGACAA 1821
Db    139 GlySerProGlyAlaProGlyThrProGlyGlyAlaLysGlnAlaGlyAspValGly--- 157
Qy   1820 TGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCAC------------ 1773
Db    158 ---SerPro------------GlyAlaProGlyThrProGlyProGlnGlyLeuProGly 172
Qy   1772 ---------TATCCAGGATGGCGAGGTCCAGGCAGATGCCCCGGC------------CCG 1734
Db    173 SerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro 192
Qy   1733 GAACCACCCTGGCCTCGGTG-------GGCTCACCCACCACCACACGTACG--------- 1690
Db    193 GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGly 212
Qy   1689 ------GAGACATCACAGGCAGAGGCCCCGCAGAGCGCGGGTGGAGGTGGGAGCAGGCCA 1636
Db    213 AlaLysGlnAlaGlyAspValGlySerProGlyAlaProGlyThrProGlyProGlnGly 232
Qy   1635 CTGCCTCCAGCACCCACGTGTCCATTAGGGAAGGGAGCTCCAGGCTTAGGGCCTGGCAGG 1576
Db    233 LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu---ProGly-Se 251
Qy   1575 AAGCTGGTCATCAGGCTGTCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGTATTTGGGC 1516
Db    251 rProGlyAlaProGlyThrProGly-----------------ProGlnGlyLeu----- 263
Qy   1515 AGGAACACCTGCTTCTCCCGGTGGTAGAGGGAGGCCA------GTGTGTAGGGCAGGATC 1462
Db    264 -------ProGlySerProGlyAlaProGlyThrProGlyGlyAlaLysGlnAlaGlyAs 281
Qy   1461 TGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTG 1402
Db    281 pValGly------------------------------------SerProGlyAlaProGl 289
Qy   1401 TGGGACAGGCATGTGGCACCGGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACT 1342
Db    289 yThrProGly--------------ProGlnGly---LeuProGlySerPro-------- 300
Qy   1341 GCTCGAGTGCCGAATCGCTGCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATG 1282
Db    301 ---------------GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPr 315
Qy   1281 GCGCACTGCAGGAACA--------------------------------GCCCCAGGCTG 1255
Db    315 oGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyTh 335
Qy   1254 CCCA---TCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTG 1198
Db    335 rProGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAlaProGlyThrProGl 355
Qy   1197 GGCACGC------CCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAG 1144
Db    355 yProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGly----- 373
Qy   1143 GTCATGAGTGCCATCCAGC------------TGCACAGCTCAGCCACGAAGAGCCGGCGC 1096
Db    374 -------LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGl 391
Qy   1095 AGGGTGCGGGGCATGCGGCAGCACAGCTGGTGCA------GCCGGGGAAGCAGGGCGCCC 1042
Db    391 ySerProGlyAlaProGlyThrProGlyGlyAlaLysGlnAlaGlyAspValGlySerPr 411
Qy   1041 AGGTTCC---GGAAAGCCAAGCGGGCCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGG 985
Db    411 oGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyTh 431
Qy    984 GCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCA---GCGCTGCCTCCTCAGCCACCAGC 928
Db    431 rPro-----------------------GlyProGlnGlyLeuProGlySerPro----- 441
Qy    927 AGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGCACTCCT-- 870
Db    442 ---------------------GlyAlaProGlyThrProGlyProGlnGlyLeuProGl 454
Qy    869 ----CCTGGGTGCCCAGGTAGGGGGCCAGGGCACTGGTGTCCCAGTCAATGGCAGGCAGG 814
Db    454 ySerProGlyAlaProGlyThr---ProGly-----------GlyAlaLysGlnAlaGl 469
Qy    813 A------GGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCC 760
Db    469 yAspValGlySerProGlyAlaPro-------------------------GlyThrPr 480
Qy    759 TGGCGACAGTGGT---CCGGGT---CCCGGAAGAGGTCAGAGAGCAGGGCCTCCAGTGGA 706
Db    480 oGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLe 500
Qy    705 GTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGCCCAGGATGAGCAGTGCCAGC 646
Db    500 uProGlySerProGly------AlaProGlyThrProGlyProGlnGlyLeuProGlySe 518
Qy    645 TCCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGA 586
Db    518 rProGlyAla------ProGlyThrProGly------GlyAlaLysGlnAlaGly---As 533
Qy    585 AAGAGGCTCAGCAGGATGCCCAAGGACAGTG---CCCAGATGAAGGCCGGCGGCGGCCA 529
Db    533 pValGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGly-------- 550
Qy    528 TAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGC 469
Db    551 ---------------------------SerProGlyAlaProGlyThrProGlyPro-- 560
Qy    468 ACTGGACCAATGCCCAGCACCATGGTCATGAACTTCTCCTCTACCCCCACTTCCAGCAGC 409
Db    561 --------------------------------GlnGlyLeuProGlySerProGlyAl 569
Qy    408 AGAGGCGGCACATAGGTGATGCCTGCGGCCAAACACACCTCCAGGCCAAAGGTTAGCAGG 349
Db    569 a-------------------------ProGlyThrProGlyProGlnGlyLeuProGl 580
Qy    348 TTGACCAGCAAGAGCTGGGCTTTCCGGTGCCGCA------GCAGGC------GGCTCACC 301
Db    580 ySerProGlyAlaProGlyThrProGlyGlyAlaLysGlnAlaGlyAspValGlySerPr 600
Qy    300 CACAGCCTCTGGACCATAGTGGGCCAGGCGGG---TAGGGCTCAGGGGGCCGTTCAGGCA 244
Db    600 oGlyAlaProGlyThr-ProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyT 620
Qy    243 CTCCAGAACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCC 184
Db    620 hrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly------- 637
Qy    183 GCCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAG 124
Db    638 --------ProGlnGlyLeuProGlySer------------ProGlyAlaProGlyThrP 651
Qy    123 CTTCTCAGCCCATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTCAT 64
Db    651 roGly--------------------------GlyAlaLysGlnAlaGlyAspValGlyS 662
Qy     63 CA---------CTCAGATCCTGGCCGA 46
```

```
Db     391 ySerProGlyAlaProGlyThrProGlyGlyAlaLysGlnAlaGlyAspValGlySerPr 411

Qy    1041 AGGTTCC---GGAAAGCCAAGCGGGCCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGG 985
           ||||:::   |||   |||          |||:::      |||
Db     411 oGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyTh 431

Qy     984 GCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCA---GCGCTGCCTCCTCAGCCACCAGC 928
            |||                       ||||||   ||||||   |||
Db     431 rPro-----------------------GlyProGlnGlyLeuProGlySerPro----- 441

Qy     927 AGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGCACTCCT-- 870
                                 |||   |||     |||   :::|||   |||
Db     442 ---------------------GlyAlaProGlyThrProGlyProGlnGlyLeuProGl 454

Qy     869 ----CCTGGGTGCCCAGGTAGGGGGCCAGGGCACTGGTGTCCCAGTCAATGGCAGGCAGG 814
               ||||||   ||||||      ||||||                ||||||||
Db     454 ySerProGlyAlaProGlyThr---ProGly-----------GlyAlaLysGlnAlaGl 469

Qy     813 A------GGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCC 760
           |       ||||||||||||:::|||                                ||
Db     469 yAspValGlySerProGlyAlaPro---------------------------GlyThrPr 480

Qy     759 TGGCGACAGTGGT---CCGGGT---CCCGGAAGAGGTCAGAGAGCAGGGCCTCCAGTGGA 706
           ||||       |||   ||||||     ||||||             ||||||
Db     480 oGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLe 500

Qy     705 GTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGCCCAGGATGAGCAGTGCCAGC 646
                 :::||||||       :::|||     |||  |||                ::
Db     500 uProGlySerProGly------AlaProGlyThrProGlyProGlnGlyLeuProGlySe 518

Qy     645 TCCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGA 586
           :|||||||||         |||||||||               ::::::   |||   ::
Db     518 rProGlyAla------ProGlyThrProGly------GlyAlaLysGlnAlaGly---As 533

Qy     585 AAGAGGCTCAGCAGGATGCCCAAGGACAGTG---CCCAGATGAAGGGCCGGCGGCGGCCA 529
           :   ||||||   |||   |||   |||      |||:::      |||
Db     533 pValGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGly-------- 550

Qy     528 TAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGC 469
                                         |||   ||||||   |||||||||
Db     551 ---------------------------SerProGlyAlaProGlyThrProGlyPro-- 560

Qy     468 ACTGGACCAATGCCCAGCACCATGGTCATGAACTTCTCCTCTACCCCCACTTCCAGCAGC 409
                                               ||||||   |||   ||
Db     561 ---------------------------------GlnGlyLeuProGlySerProGlyAl 569

Qy     408 AGAGGCGGCACATAGGTGATGCCTGCGGCCAAACACACCTCCAGGCCAAAGGTTAGCAGG 349
           |                           |||   ||||||   |||   |||   ||
Db     569 a---------------------------ProGlyThrProGlyProGlnGlyLeuProGl 580

Qy     348 TTGACCAGCAAGAGCTGGGCTTTCCGGTGCCGCA-----GCAGGC------GGCTCACC 301
           |   |||         |||         |||   |||      |||      ||||||||
Db     580 ySerProGlyAlaProGlyThrProGlyGlyAlaLysGlnAlaGlyAspValGlySerPr 600

Qy     300 CACAGCCTCTGGACCATAGTGGGCCAGGCGGG---TAGGGCTCAGGGGGCCGTTCAGGCA 244
           |   |||   |||         ||||||   |||      ||||||   |||      ||||
Db     600 oGlyAlaProGlyThr-ProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyT 620

Qy     243 CTCCAGAACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCC 184
           |||||          |||         :::|||   |||         :::|||
Db     620 hrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly------- 637

Qy     183 GCCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAG 124
                   |||:::      |||||||:::            |||         |||
Db     638 --------ProGlnGlyLeuProGlySer------------ProGlyAlaProGlyThrP 651

Qy     123 CTTCTCAGCCCATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTCAT 64
                                              ||||||   |||   |||   |||   |
Db     651 roGly---------------------------GlyAlaLysGlnAlaGlyAspValGlyS 662

Qy      63 CA---------CTCAGATCCTGGCCGA 46
           ||         ::: |||||||||||
Db     662 erProGlyAlaMet-AspProGlyArg 670
```

```
RESULT 21
US-08-397-633A-36
; Sequence 36, Application US/08397633A
; Patent No. 5773577
;  GENERAL INFORMATION:
;    APPLICANT:  Cappello, Joseph
;    TITLE OF INVENTION:  PRODUCTS COMPRISING SUBSTRATESCAPABLE
;    TITLE OF INVENTION:  OF ENZYMATIC CROSS-LINKING
;    NUMBER OF SEQUENCES:  105
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/397,633A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rowland, Bertram I
;      REGISTRATION NUMBER:  20,015
;      REFERENCE/DOCKET NUMBER:  A-58848-1/BIR PROP-011-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 781-1989
;      TELEFAX:  (415) 398-3249
;      TELEX:  910 277299
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  682 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-397-633A-36

Alignment Scores:
Pred. No.:               2.04e-20      Length:          682
Score:                   370.50        Matches:         249
Percent Similarity:      35.26%        Conservative:    44
Best Local Similarity:   29.96%        Mismatches:      294
Query Match:             5.95%         Indels:          246
DB:                      1             Gaps:            56

US-09-759-143-110 (1-3410) x US-08-397-633A-36 (1-682)

Qy    2314 AAAAACCCTTCTCTAGGTGTGTCTCAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGT 2255
             :::||||||       ||||||:::||||||                  |||:::
Db      12 GluAsnPro------GlyValThrGlnLeuAsn------------ArgLeuAlaAlaHis 25

Qy    2254 AATCCACCTGCAGAGTCCCCGCATTCCAGTGCATGGAGCCCTTCT--------------- 2210
              |||   |||   |||   |||         :::      :::|||
Db      26 ProProPheAlaSerAspProMetGlyAlaProGlyThrProGlyProGlnGlyLeuPro 45

Qy    2209 GGCCTCCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCT 2150
            |||   |||         |||   |||   :::||||||         |||||||||
Db      46 GlySerProGlyAlaProGlyThrProGlyPro---GlnGlyLeu---ProGlySerPro 63

Qy    2149 AGAGACTGGGGAGAGAGGAGAGGGACGCCCCAGCCCCCAGCTGTGCAGCTACGCACCTCA 2090
                                ||||||||   |||         :::
Db      64 ------------GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGly 79

Qy    2089 GCA------GCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGCAA---------- 2046
            |||            ||||||:::::::::            ||||||
Db      80 AlaProGlyThrProGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAlaPro 99
```

```
US-08-642-255-126
; Sequence 126, Application US/08642255
; Patent No. 5773249
;  GENERAL INFORMATION:
;    APPLICANT:  CAPPELLO, Joseph
;    APPLICANT:  FERRARI, Franco A.
;    TITLE OF INVENTION:  High Molecular Weight Collagen-Like
;    TITLE OF INVENTION:  Protein Polymers
;    NUMBER OF SEQUENCES:  135
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/642,255
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROWLAND, Bertram I.
;      REGISTRATION NUMBER:  20,015
;      REFERENCE/DOCKET NUMBER:  A55556-3/BIR
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 494-8700
;      TELEFAX:  (415) 494-8771
;      TELEX:  910 277299 FHT UR
;  INFORMATION FOR SEQ ID NO:  126:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  682 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-642-255-126

Alignment Scores:
Pred. No.:               2.04e-20      Length:        682
Score:                   370.50        Matches:       249
Percent Similarity:      35.26%        Conservative:  44
Best Local Similarity:   29.96%        Mismatches:    294
Query Match:             5.95%         Indels:        246
DB:                      1             Gaps:          56

US-09-759-143-110 (1-3410) x US-08-642-255-126 (1-682)

Qy       2314 AAAAACCCTTCTCTAGGTGTGTCTCAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGT 2255
                :::|||||        ||||||:::||||||                   |||:::
Db         12 GluAsnPro------GlyValThrGlnLeuAsn-----------ArgLeuAlaAlaHis 25

Qy       2254 AATCCACCTGCAGAGTCCCCGCATTCCAGTGCATGGAGCCCTTCT-------------- 2210
                |||   |||       |||      :::       :::|||
Db         26 ProProPheAlaSerAspProMetGlyAlaProGlyThrProGlyProGlnGlyLeuPro 45

Qy       2209 GGCCTCCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCT 2150
                |||   |||          |||      |||    :::||||||      |||||||||
Db         46 GlySerProGlyAlaProGlyThrProGlyPro---GlnGlyLeu---ProGlySerPro 63

Qy       2149 AGAGACTGGGGAGAGAGGAGAGGGACGCCCCAGCCCCCAGCTGTGCAGCTACGCACCTCA 2090
                                    |||||||||   |||        :::
Db         64 -----------GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGly 79

Qy       2089 GCA------GCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGCAA---------- 2046
                |||              ||||||:::::::::            ||||||
Db         80 AlaProGlyThrProGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAlaPro 99

Qy       2045 ---------------CAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAACAGGAG 2001
                                |||   |||         ||||||    ||||||
Db        100 GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro 119

Qy       2000 CGGGGAGCTGGGACCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTTCTACG 1941
                :::|||                      |||   |||   ||||||
Db        120 GlnGlyLeuProGlySerProGlyAlaProGlyThrPro-------------------- 132

Qy       1940 CTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGA 1881
                                         |||   ||||||
Db        133 --------------------GlyProGlnGlyLeuPro-------------------- 138

Qy       1880 CCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTGGACAA 1821
                   ||||||         ||||||              |||   |||      |||
Db        139 GlySerProGlyAlaProGlyThrProGlyGlyAlaLysGlnAlaGlyAspValGly--- 157

Qy       1820 TGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCAC------------ 1773
                   ||||||          |||   |||||||||   |||
Db        158 ---SerPro------------GlyAlaProGlyThrProGlyProGlnGlyLeuProGly 172

Qy       1772 ---------TATCCAGGATGGCGAGGTCCAGGCAGATGCCCCGGC-----------CCG 1734
                         ||||||        ||||||            ||||||         |||
Db        173 SerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro 192

Qy       1733 GAACCACCCTGGCCTCGGTG-------GGCTCACCCACCACCACACGTACG--------- 1690
                      |||   |||:::          |||||||||         |||
Db        193 GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGly 212

Qy       1689 ------GAGACATCACAGGCAGAGGCCCCGCAGAGCGCGGGTGGAGGTGGGAGCAGGCCA 1636
                    :::               :::|||   :::   |||      |||   :::
Db        213 AlaLysGlnAlaGlyAspValGlySerProGlyAlaProGlyThrProGlyProGlnGly 232

Qy       1635 CTGCCTCCAGCACCCACGTGTCCATTAGGGAAGGGAGCTCCAGGCTTAGGGCCTGGCAGG 1576
                ||||||   :::|||      |||         |||      ||||||   ||||||
Db        233 LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeu---ProGly-Se 251

Qy       1575 AAGCTGGTCATCAGGCTGTCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGTATTTGGGC 1516
                   :::   |||   |||                           |||   |||:::
Db        251 rProGlyAlaProGlyThrProGly-----------------ProGlnGlyLeu----- 263

Qy       1515 AGGAACACCTGCTTCTCCCGGTGGTAGAGGGAGGCCA------GTGTGTAGGGCAGGATC 1462
                   |||   |||||||||      |||   |||                :::||||||
Db        264 -------ProGlySerProGlyAlaProGlyThrProGlyGlyAlaLysGlnAlaGlyAs 281

Qy       1461 TGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTG 1402
                 |||                                         ||||||
Db        281 pValGly-----------------------------------SerProGlyAlaProGl 289

Qy       1401 TGGGACAGGCATGTGGCACCGGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACT 1342
                 |||                  |||||||||   ||||||      |||
Db        289 yThrProGly--------------ProGlnGly---LeuProGlySerPro-------- 300

Qy       1341 GCTCGAGTGCCGAATCGCTGCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATG 1282
                                  ||||||      |||   |||:::         ||||||
Db        301 ----------------GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPr 315

Qy       1281 GCGCACTGCAGGAACA-------------------------------GCCCCAGGCTG 1255
                      ||||||                                   |||||||||
Db        315 oGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyTh 335

Qy       1254 CCCA---TCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTG 1198
                 |||         :::                  |||      |||      |||
Db        335 rProGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAlaProGlyThrProGl 355

Qy       1197 GGCACGC------CCTGGTACAGCCCCTCGCCCACGAATCCGTGTAAAACAGCGTGAAG 1144
                 :::            ||||||:::            |||      |||
Db        355 yProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGly----- 373

Qy       1143 GTCATGAGTGCCATCCAGC-----------TGCACAGCTCAGCCACGAAGAGCCGGCGC 1096
                    :::|||   |||                  |||         |||:::
Db        374 -------LeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGl 391

Qy       1095 AGGGTGCGGGGCATGCGGCAGCACAGCTGGTGCA------GCCGGGGAAGCAGGGCGCCC 1042
                   ||||||   |||:::         ||||||            ||||||:::   |||   ||
```

```
Qy   1879 CAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTGGACAAT 1820
            ||||||||          |||     |||     |||     |||
Db    184 --AspProGly---------ProProGlyAlaGlnGlyProAlaGlyProGlyGlySerA 200

Qy   1819 GGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCACTATCCAGGATGGC 1760
              |||             |||||||||||       |||          |||
Db    200 rgAspPro------------GlyProProGlyAlaGlnGly---------ProAlaGlyP 213

Qy   1759 GAGGTCCAGGCAGATGCCCCGGCCCGGAA------------------------------- 1731
            |||     |||  |||||||||
Db    213 roGlyGlySerArgAspProGlyProProGlyAlaHisGlyProAlaGlyProLysGlyA 233

Qy   1730 -------CCACCCTGGCCTCGGTGGGCTCACCCACCACCACACGTACGGAGACATCACA 1679
                   |||      |||:::  ||||||  |||          :::      |||
Db    233 laHisGlyProAlaGlyProLysGlyAlaHisGlyProAlaGlyProLysGlyAlaHisG 253

Qy   1678 G------------------GCAGAGGCCCCGCAGAGCGCGGGTGGAGGTGGGAGCAGG- 1639
                                |||:::  |||    ||||||  |||||||||||
Db    253 lyProAlaGlyProLysGlyAlaGlnGlyPro------AlaGlyProGlyGlySerArgA 271

Qy   1638 --CCACTGCCTCCAGCACCCACGTGTCCATTAGGGAAGGGAGCT--------------- 1597
            |||   ||||||          |||    |||    |||
Db    271 spProGlyProProGlyAlaGlnGlyProAlaGlyProGlyGlySerArgAspProGlyP 291

Qy   1596 --CCAGGC------------TTAGGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCAC 1551
            ||||||            ||||||||| ||||||            |||     |||
Db    291 roProGlyAlaGlnGlyProAlaGlyProGly-GlySerArgAspProGlyProPro--- 309

Qy   1550 TGCTAGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGT 1491
                                         |||       ||||||    |||||||||
Db    310 ------------------------------GlyAlaGlnGlyProAlaGlyProGlyGly 319

Qy   1490 AGA---------GGGAGGCCAGTGTGTAGG------GCAGGATCTGCAGGGCTGAGAAGG 1446
                        |||   |||      :::      ||||||     |||    |||
Db    320 SerArgAspProGlyProProGlyAlaGlnGlyProAlaGlyProGlyGlySerArg--- 338

Qy   1445 TGAACCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGG 1386
                                         |||   |||            |||
Db    339 ---------------------------AspProGlyProProGlyAlaGlnGly------ 347

Qy   1385 CACCGGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATC 1326

Db    347 ------------------------------------------------------------ 347

Qy   1325 GCTGCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACA 1266
                ||||||||||||        |||   ||||||               |||
Db    348 ------ProAlaGlyProGlyGlySerArgAspProGlyProProGlyAlaGlnGlyPro 365

Qy   1265 GCC--------------------CCAGGCTGCCCATCCGAACGCCTTCATCATAGTGTC 1227
            |||                    ||||||   |||      :::
Db    366 AlaGlyProGlyGlySerArgAspProGlyProProGlyAlaGlnGlyProAlaGlyPro 385

Qy   1226 TCCGGGCCTCGGTGCCCG-----------GCTCAGCTCTGGGCACGCCCTGGTACAGCC 1179
               |||   |||   |||           ||||||     |||   ||||||   :::
Db    386 GlyGlySerArgAspProGlyProProGlyAlaGlnGlyProAlaGlyProGlyGlySer 405

Qy   1178 CCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGCTGCACA 1119
               |||     |||
Db    406 ArgAspProGlyProPro------------------------------------------ 411

Qy   1118 GCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACAGCT---GGTGCA 1062
                                       ||||||   |||            |||
Db    412 ------------------------------GlyAlaGlnGlyProAlaGlyProGlyGly 421

Qy   1061 GCCGGGGAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGGCCCGGCATGGACAGCAGT 1002
            :::   :::   |||   ||||||:::      |||:::|||||||||
Db    422 SerArgAspProGlyProProGlyAlaGlnGlyProAlaGlyProGly------------ 437

Qy   1001 GGGGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCT 942
            |||:::         ||||||                      |||||||||   |||
Db    438 GlySerArgAspProGlyProPro------------GlyAlaGlnGlyProAlaGlyPro 453

Qy    941 ---------------CCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGA 897
                         |||  ||||||                            |||
Db    454 GlyGlySerArgAspProGlyProPro-----------------------GlyAlaHis 465

Qy    896 GGGTGAGCAGGCCAAAGAGGCACT---------------CCTCCTGGTGCCCAGGT--- 855
            |||  ||||||    :::|||                   |||    |||    |||
Db    466 GlyProAlaGlyProLysGlyAlaHisGlyProAlaGlyProLysGlyAlaHisGlyPro 485

Qy    854 AGGGGGCCA------------GGGCACTGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGC 807
               ||||||            |||           |||         |||
Db    486 AlaGlyProLysGlyAlaHisGlyProAlaGlyProLysGlyAlaGlnGlyProAlaGly 505

Qy    806 CCAGGCAGCCCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGT 747
            ||||||      :::|||                          ||||||    |||
Db    506 ProGlyGlySerArgAsp------------------ProGlyProProGlyAlaGlnGly 519

Qy    746 CCG---GGTCCCGGAAGAGGTCAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGC 690
            |||   |||||||||       :::   ||||||                  ||||||
Db    520 ProAlaGlyProGlyGlySerArgAspProGlyPro------------------ProGly 533

Qy    689 CACAGAAGTCCAGCAGCCCCA----------------------CGCCCAGGATGAGCAGTG 651
              :::   ||||||   |||                           ||||||
Db    534 AlaGlnGlyProAlaGlyProGlyGlySerArgAspProGlyProProGlyAlaGlnGly 553

Qy    650 CCAGCT------CCAGGGGCCTGGGATCCGGGC---ACAGCAGCCCTGCTAGCCAGCCGG 600
            ||||||            |||:::   |||||||||          |||    ||||||||
Db    554 ProAlaGlyProGlyGlySerArgAspProGlyProProGlyAlaGln-GlyProAlaGl 573

Qy    599 CCCTTGGGATGAGAAAGAG---GCTCAGCAGGATGCCCAAGGACAGTGCCCAGATGAAGG 543
            ||||            ::::::                |||||||||   |||       |
Db    573 yProGlyGlySerArgAspProGlyProProGlyAlaGlnGly-ProAlaGlyProGlyG 593

Qy    542 GCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTAGGAGCGGGACAC 483
            ||   |||   |||       |||                  |||   |||  |||
Db    593 lySerArgAspProGlyProProGlyAlaGlnGlyProAlaGlyPro-GlyGlySerArg 612

Qy    482 AGACCAGGCCCAGCACTGGACCAATGCCCAGCA---CCATGGTCATGAACTTCTCCTCTA 426
               |||||||||         |||   ||||||   |||                |||
Db    613 AspProGlyProProGlyAlaGlnGlyProAlaGlyProGlyGlySerArgAspProGly 632

Qy    425 CCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCCAAACACACCTCCA 366
            |||      |||   |||:::                        |||       |||
Db    633 Pro------ProGlyAlaGlnGly---------------------ProAlaGlyProGly 643

Qy    365 GGCCAAAGGTTAGCAGGTTGACCAGCAAGA---------GCTGGGCTTTCCGGTGCCGCA 315
            |||   |||      |||   |||                  ||||||     |||:::
Db    644 GlySerArgAspProGlyProProGlyAlaGlnGlyProAlaGlyProGlyGlySerArg 663

Qy    314 GCAGGCGGCTCACCCACAGCCTCTGGACCATAGTGGGCCAGGCGGGTAGGGCTCAGGGGG 255
               |||   |||   |||   ||||||  :::||||||||||||                ||
Db    664 AspProGlyProProGlyAlaGlnGlyPro-AlaGlyProGlyGlySerArgAspProGl 683

Qy    254 CCGTTCAGGCACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCC 195
            |      |||                      |||   |||              :::
Db    683 yProProGlyAlaGln---------------GlyProAlaGlyProGlyGlySerArgAs 698

Qy    194 TCCTTGCTGCCGCCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGG 135
            |||                          |||||||||||          |||    ||
Db    698 pPro--------------------GlyProProGlyAlaHis------GlyProAlaGl 709

Qy    134 TGCCGGTCCAGCTTCTCAGCCCATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTGGGG 75
            | |||                  |||                  |||    |||
Db    709 y-ProLysGlyAlaHisGlyProAlaGlyProLysGlyAlaHisGlyProAlaGlyProL 729

Qy     74 ACACGTCTCATCACTCAGATCCTGGCCGAGGCGCGCGG------CTGTCACCCGGAGCCA 21
              :::|||                   :::||||||:::              ||||||   |
Db    729 ysGlyAlaHisGlyProAlaGlyProLysGlyAlaGlnGlyProAlaGlyProGlyGlyS 749

Qy     20 GC 19
          ||
Db    749 er 749
```

RESULT 20

```
                 ||||||                     |||||| |||  |||||| |||||||
Db        334 roProGlySerArgAspProGlyProProGlyAlaProGlyProAlaGlyProProGlyS 354

Qy       1036 CCGGAAAGCCAAGCGGGCCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAG 977
              ||   |||   ||||||
Db        354 erArgAspProGlyProProGlyAlaPro------------------------------ 363

Qy        976 CCCTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCT---------------CCTCAGCCAC 932
                                 ||||||||  |||            |||  ||||
Db        364 -------------------GlyProAlaGlyProProGlySerArgAspProGlyProP 377

Qy        931 CAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGCACT- 873
              ||                     |||   |||  ||||||   |||:::
Db        377 ro-----------------------GlyAlaProGlyProAlaGlyProProGlySerA 389

Qy        872 -----------CCTCCTGGGTGCCCAGGTAGGGGG------CCAGGGCACTGGTGTCCCA 830
                         |||||||||  ||||||          ||||||      |||
Db        389 rgAspProGlyProProGlyAlaProGlyProAlaGlyProProGlySerArgAspPro- 408

Qy        829 GTCAATGGCAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCATAGAC 770
                               |||   |||||:::|||
Db        409 -----------------GlyProProGlyAlaProGlyPro-----------------A 417

Qy        769 AGAGTAGGCCTGGCGACAGTGGTCCGGGT---CCCGGAAGAGGTCAGAGAGCAGGGCCTC 713
                         ||||||        ||||||  ||||||        |||||||||
Db        417 laGlyProProGlySerArgAspProGlyProProGlyAlaProGlyProAlaGlyPro- 436

Qy        712 CAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCA---CGCCCAGGATGAG 656
                               ||||||  |||  |||     |||   ||||||   |
Db        437 -----------------ProGlySerArgAspProGlyProProGlyAlaProGlyProA 451

Qy        655 CAGTGCCAGCTCCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGGCCCT 596
              ||  |||  ||||||:::  ||||||||            |||       |||
Db        451 laGlyPro---ProGlySerArgAspProGlyProProGly------AlaHisGlyProA 468

Qy        595 TGGGATGAGAA------AGAGGCTCAGCAGGATGCCCAAGGACAGTGCCCAGATGAAGGG 542
                |||  :::          |||  ||||||               |||:::
Db        468 laGlyProLysGlyAlaHisGlyProAlaGly---------------ProLysGlyAlaH 483

Qy        541 CCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTAGGAGCGGGACACA 482
                |||                    |||  |||              ||||||
Db        483 isGlyProAlaGlyProLysGlyAlaHisGlyProAlaGlyProLysGlyAlaProGlyP 503

Qy        481 GACCAGGCCCAGCACTGGACCAATGCCCAGCACCATGGTCATGAACTTCTCCTCTACCCC 422
                  ||||||          :::  |||  |||            |||    |
Db        503 roAlaGlyProProGlySerArgAspProGlyProProGlyAlaProGlyProAlaGlyP 523

Qy        421 CACTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCCAAACACACCTCCAGGCC 362
              ||     :::   :::                        |||   |||||||||
Db        523 roProGlySerArgAspProGlyProProGlyAlaProGlyProAlaGlyProProGlyS 543

Qy        361 AAAGG---------TTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGTGCCGCAGCAG 311
                |||                 |||  |||    ||||||   |||:::
Db        543 erArgAspProGlyProProGlyAlaProGlyProAlaGlyProProGlySerArgAspP 563

Qy        310 GCGGCTCACCCACAGCCTCTGGACCATAGTGGGCCAGGCGGGTAGGGCTCAGGGGGCCGT 251
                |||  |||  |||  ||||||  :::||||||  |||               |||
Db        563 roGlyProProGlyAlaProGlyPro-AlaGlyProProGlySerArgAspProGlyPro 582

Qy        250 TCAGGCACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCT 191
                 |||  |||                            |||   ||||||
Db        583 ProGlyAlaPro----------------------------GlyProAlaGlyProPro 592

Qy        190 TGCTGCCGCCAACTG--------------------CCTAGGAATCAGCCAGGCGCCCAT 152
                   |||                           |||      ||||||:::
Db        593 GlySerArgAspProGlyProProGlyAlaProGlyProAlaGlyProProGlySerArg 612

Qy        151 TTCTGCCAGCCCTTTGGTGCC---GGTCCAGCTTCTCAGCCC 113
                       |||  ||||||  |||||||||     |||
Db        613 AspProGlyProProGlyAlaProGlyProAlaGlyProPro 626
```

RESULT 19
US-09-219-849-6

```
; Sequence 6, Application US/09219849
; Patent No. 6150081
; GENERAL INFORMATION:
;  APPLICANT:  VAN HEERDE, GEORGE V.
;  APPLICANT:  VAN RIJN, ALEXIS C.
;  APPLICANT:  BOUWSTRA, JAN B.
;  APPLICANT:  DE WOLF, FREDERIK A.
;  APPLICANT:  MOOBROEK, ANDREAS
;  APPLICANT:  WERTEN, MARC W.T.
;  APPLICANT:  WIND, RICHELE D.
;  APPLICANT:  VAN DEN BOSCH, TANJA J.
;  TITLE OF INVENTION:  SILVER HALIDE EMULSIONS WITH RECOMBINANT COLLAGEN
;  TITLE OF INVENTION:  SUITABLE FOR PHOTOGRAPHIC APPLICATION AND ALSO THE
;  TITLE OF INVENTION:  PREPARATION THEREOF
;  FILE REFERENCE:  2728-2
;  CURRENT APPLICATION NUMBER: US/09/219,849
;  CURRENT FILING DATE:  1998-12-23
;  NUMBER OF SEQ ID NOS:  50
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 960
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Illustrative
;   OTHER INFORMATION: amino acid sequence
US-09-219-849-6

Alignment Scores:
Pred. No.:              2.11e-20      Length:         960
Score:                  371.00        Matches:        255
Percent Similarity:     32.39%        Conservative:   32
Best Local Similarity:  28.78%        Mismatches:     301
Query Match:            5.96%         Indels:         299
DB:                     4             Gaps:           53

US-09-759-143-110 (1-3410) x US-09-219-849-6 (1-960)

Qy       2359 GGGCTTAGAGAT---GGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCT 2303
                |||  ||||||  |||  ||||||                  |||    |||
Db         62 GlySerArgAspProGlyProProGly---------AlaGlnGlyProAlaGlyPro--- 77

Qy       2302 CTAGGTGTGTCTCAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCA 2243
                                 ||||||          :::       |||  |||  |||
Db         78 -----------------GlyGly---------SerArgAspProGlyProProGlyAla 88

Qy       2242 GAGTCCCCGCATTCC-----------------------------AGTGCATGGAGCCCT 2213
              :::   |||                                          |||   |||
Db         89 GlnGlyProAlaGlyProGlyGlySerArgAspProGlyProProGlyAlaGlnGlyPro 108

Qy       2212 TCTGGCCTCCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGCCTCCAGTCAGGCAGC 2153
              :::|||                                                 |||
Db        109 AlaGly---------------------------------------------ProGlyGly 113

Qy       2152 CCTAGAGACTGGGGAGAGAGGAGAGGGACGCCCCAGCCCCCAGCTGTGCAGCTACGCACC 2093
                ||||||                        |||   ||||||     |||
Db        114 SerArgAsp--------------------ProGlyProProGlyAlaGln------Gly 124

Qy       2092 TCAGCAGCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGG 2033
                 |||       ||||||||||||:::|||                |||:::
Db        125 ProAlaGlyProGlyGlySerArgAspProGlyProProGlyAlaGlnGlyProAlaGly 144

Qy       2032 CCA------GCCCGGCAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGACCCAGT--- 1982
              |||      :::|||  |||                 |||      ||||||
Db        145 ProGlyGlySerArgAspProGlyProProGlyAlaGlnGlyProAlaGlyProGlyGly 164

Qy       1981 ------GAGGCAGGCCCTCCA-----------CCCCAATGTGCTGGAAGTTTTCTACGC 1940
                    :::   |||||||||               |||         |||
Db        165 SerArgAspProGlyProProGlyAlaGlnGlyProAlaGlyProGly----------- 180

Qy       1939 TGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGAC 1880
                       |||  ||||||
Db        181 ---------Gly--SerArg--------------------------------------- 183
```

```
Qy        294 CTCTGGACCATAGTGGGCCAGGCGGG---TAGGGCTCAGGGGGCCGTTCAGGCACTCCAG 238
              |  |||      ||||||  |||      ||||||   |||      |||||||||
Db        674 aProGlyThr-ProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProG 694

Qy        237 AACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAAC 178
                          |||      :::|||   |||       :::|||
Db        694 lyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly------------ 709

Qy        177 TGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGCTTCTC 118
                 |||:::      ||||||:::          |||        |||          |
Db        710 --ProGlnGlyLeuProGlySerProGlyAla---ProGlyThrProGlyGluGlyGlnG 728

Qy        117 AGCCCATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTCATCA---- 62
              ||       |||                ||||||   |||   |||   |||   |||
Db        728 lnHisHisLeuGly-----------GlyAlaLysGlnAlaGlyAspValGlySerProG 744

Qy         61 -----CTCAGATCCTGGCCGA 46
                   ::: ||||||||||||
Db        744 lyAlaMet-AspProGlyArg 750
```

RESULT 18
US-09-219-849-5
; Sequence 5, Application US/09219849
; Patent No. 6150081
; GENERAL INFORMATION:
; APPLICANT: VAN HEERDE, GEORGE V.
; APPLICANT: VAN RIJN, ALEXIS C.
; APPLICANT: BOUWSTRA, JAN B.
; APPLICANT: DE WOLF, FREDERIK A.
; APPLICANT: MOOBROEK, ANDREAS
; APPLICANT: WERTEN, MARC W.T.
; APPLICANT: WIND, RICHELE D.
; APPLICANT: VAN DEN BOSCH, TANJA J.
; TITLE OF INVENTION: SILVER HALIDE EMULSIONS WITH RECOMBINANT COLLAGEN
; TITLE OF INVENTION: SUITABLE FOR PHOTOGRAPHIC APPLICATION AND ALSO THE
; TITLE OF INVENTION: PREPARATION THEREOF
; FILE REFERENCE: 2728-2
; CURRENT APPLICATION NUMBER: US/09/219,849
; CURRENT FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 960
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Illustrative
;   OTHER INFORMATION: amino acid sequence
US-09-219-849-5

```
Alignment Scores:
Pred. No.:                1.14e-20      Length:         960
Score:                    374.50        Matches:        237
Percent Similarity:       32.90%        Conservative:   19
Best Local Similarity:    30.46%        Mismatches:     293
Query Match:              6.02%         Indels:         229
DB:                       4             Gaps:           44
```

US-09-759-143-110 (1-3410) x US-09-219-849-5 (1-960)

```
Qy       2257 GGTAATCCA---CCTGCAGAGTCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCT 2201
              |||   |||   ||||||      |||   |||      |||   |||   |||   |||
Db         19 GlyAlaProGlyProAlaGlyProProGlySerArgAspProGlyProProGlyAlaPro 38

Qy       2200 GTATAAGTCCAGACTGAAACCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCT------AGA 2147
                                                  :::|||   |||      |||
Db         39 Gly-------------------------------ProAlaGlyProProGlySerArg 47

Qy       2146 GACTGGGGAGAGAGGAGAGGGACGCCCCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCA 2087
              |||   |||         |||   |||   |||                      |||
Db         48 AspProGlyPro---ProGlyAlaProGlyPro-----------------------Ala 58

Qy       2086 GCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCC 2027
                        ||||||||:::|||         |||:::            ||||||
Db         59 GlyProProGlySerArgAspProGlyProProGlyAlaPro---------GlyProAla 75

Qy       2026 CGGCAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGGA---CCCAGTGAGGCAGGCCCT 1970
                   |||   |||:::      |||      |||   |||      |||||||||
Db         76 GlyProPro---GlySerArgAspProGlyProProGlyAlaProGlyProAlaGlyPro 94

Qy       1969 CCACCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAA 1910
              |||                                   |||  ||||||
Db         95 Pro--------------------------------Gly--SerArg---------- 98

Qy       1909 TACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCCTGCGGGAGACACCATATA 1850
                                            |||||||||         |||
Db         99 ------------------------------AspProGly---------ProProGlyA 105

Qy       1849 GGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCATAAACAGGGATGGGGCCACCTG 1790
                       |||      |||            |||            |||||||||
Db        105 laProGlyProAlaGlyProProGlySerArgAspPro-----------GlyProProG 121

Qy       1789 GGACAGCAGGA---AGGCACTATCCAGGATGGCGAGGTCCAGGC--------------- 1749
              ||      |||         ||||||   |||   ||||||
Db        121 lyAlaProGlyProAlaGlyProProGlySerArgAspProGlyProProGlyAlaProG 141

Qy       1748 ----------------------AGATGCCCCGGCCCGGAACCACCCTGGCCTCGGTGGG 1712
                                    |||   |||||||||   |||      |||     |
Db        141 lyProAlaGlyProProGlySerArgAspProGlyPro---ProGlyAlaProGlyProA 160

Qy       1711 CTCACCCACCACCACACGTACGGAGACATCACAGGCAGAGGCCCCGCAGAGCGCGGGTGG 1652
              ||   ||||||                :::               ||||||   ||||||
Db        160 laGlyProPro--GlySerArgAspProGlyProProGlyAlaProGlyProAlaGlyPr 179

Qy       1651 AGGTGGGAGCAGG---CCACTGCCTCCAGCACCCACGTGTCCATTAGGG---------- 1606
                 |||||||||   |||   ||||||            |||   |||
Db        179 oProGlySerArgAspProGlyProProGlyAlaProGlyProAlaGlyProProGlySe 199

Qy       1605 ----------------AAGGGAGCTCCAGGC---TTAGGGCCTGGCAGGAAGCTGGTCAT 1565
                                ||||||||||||      ||||||    ||||||
Db        199 rArgAspProGlyProProGlyAlaProGlyProAlaGlyPro-ProGlySerArgAspP 219

Qy       1564 CAGGCTGTCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTG 1505
                 |||   |||                  |||   |||      |||     ||||
Db        219 roGlyProProGlyAlaHisGlyProAlaGlyProLysGly------AlaHisGlyProA 237

Qy       1504 CTTCTCCCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGT 1445
              ||   |||   |||            |||               |||:::   |||
Db        237 laGlyProLysGlyAlaHisGlyProAlaGlyProLysGlyAlaHisGlyPro------- 254

Qy       1444 GAACCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGC 1385
                                |||   :::|||   |||            |||
Db        255 -----------AlaGlyProLysGlyAlaProGlyProAlaGlyProProGlySerArgA 271

Qy       1384 ACCGGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCG 1325
                       |||   |||
Db        271 spProGlyProProGly--------------------------------------AlaP 278

Qy       1324 CTGCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACAG 1265
                    ||||||||||||         |||   ||||||
Db        278 roGlyProAlaGlyProProGlySerArgAspProGly--------------------P 291

Qy       1264 CCCCAGGCTGCCCATCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTC 1205
                 ||||||   |||                     :::||||||         |||
Db        291 roProGlyAlaProGlyPro-----------------AlaGlyPro------ProGlyS 303

Qy       1204 AGCTCTGGGCACGCCCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAA 1145
                          |||      ||||||   |||      |||
Db        303 erArgAspProGlyProProGlyAlaProGlyProAlaGlyPro--------------- 317

Qy       1144 GGTCATGAGTGCCATCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGG 1085
                          |||   |||            |||      |||            |||
Db        318 ----------ProGlySerArgAspProGlyProProGlyAlaProGlyProAlaGlyP 334

Qy       1084 CATGCGGCAGC------------ACAGCTGGTGCAGCCGGGGAAGCAGGGCGCCCAGGTT 1037
```

```
                 |||||||||                       :::|||:::|||    |||
Db        91 HisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAlaProGlyThrPro 110

Qy      2239 TCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACC 2180
              |||                |||||||   |||   |||          |||
Db       111 GlyProGlnGlyLeuProGlySerPro---GlyAlaPro---------GlyThrProGly 126

Qy      2179 CCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCCC 2120
              |||    ::::||||||     |||||||||                  |||||||||
Db       127 Pro---GlnGlyLeu---ProGlySerPro-----------GlyAlaProGlyThrPro 140

Qy      2119 CAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGT------GGCAGCAGAGAG 2066
                |||       :::        :::     |||   |||    |||       :::
Db       141 GlyProGlnGlyLeuPro---GlySerProGlyAlaProGlyThrProGlyGluGlyGln 159

Qy      2065 CCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAAC 2006
                |||:::   |||                                   ||||||
Db       160 GlnHisHisLeuGly-----------------------------------GlyAlaLys 167

Qy      2005 AGGAGCGGGGAGCTGGGACCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTT 1946
             ::::::|||::::::|||        |||   |||   ||||||
Db       168 GlnAlaGlyAspValGlySerPro--GlyAlaProGlyThrPro--------------- 181

Qy      1945 CTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAAT 1886
                                                |||   ||||||
Db       182 -------------------------GlyProGlnGlyLeuPro----------------- 187

Qy      1885 GGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTG 1826
                            ||||||       ||||||      |||
Db       188 -----GlySerProGlyAlaProGlyThrProGlyProGln------------------ 199

Qy      1825 GACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCACTATCCAG 1766
                                          |||   ||||||:::    |||         ||||
Db       200 -------------------------GlyLeuProGlySerProGlyAla------ProG 209

Qy      1765 GATGGCGAGGTCCAGGCAGATGCCCCGGC-----------CCGGAACCACCCTGGCCTC 1718
             ||      ||||||        ||||||                 |||     ||||||
Db       209 lyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro-GlyThrProGlyGlu 228

Qy      1717 GGTGGGCTCACCCACCACCACACGTACGGAGACATCACAGGCA---------GAGGCCCC 1667
              |||            ||||||     |||       ||||||             :::||
Db       229 Gly------GlnGlnHisHisLeuGlyGly-AlaLysGlnAlaGlyAspValGlySerPr 246

Qy      1666 GCAGAGCGCGGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGG 1607
             |    :::     |||      |||     :::     ||||||     :::|||
Db       246 oGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro------------- 261

Qy      1606 GAAGGGAGCTCCAGGCTTAGGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCT 1547
                ||||||||||||        ||||||    :::     :::     |||   |||
Db       262 ----GlyAlaProGlyThr---ProGlyProGlnGlyLeu-ProGlySerPro------- 276

Qy      1546 AGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGTAGAG 1487
                                            |||   |||   :::||||||    :::|
Db       277 ------------------------------GlyAlaProGlyThrProGlyProGlnG 286

Qy      1486 GGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGC 1427
             ||    |||               ||||||   |||       |||   ||||||||||||
Db       286 lyLeuPro-----------GlySerProGlyAla-ProGlyThrProGlyGluGly--- 300

Qy      1426 TGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACCGGCAG---------- 1377
                        ||||||                  ||||||         |||
Db       301 ---GlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAla 319

Qy      1376 --------------CCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGC 1332
                            |||||||||   ||||||       |||
Db       320 ProGlyThrProGlyProGlnGly---LeuProGlySerPro------------------ 332

Qy      1331 CGAATCGCTGCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCA 1272
                       ||||||   |||   |||:::        ||||||
Db       333 ------GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro 350

Qy      1271 GGAACA-------------------------------GCCCCAGGCTGCCCATCCGAA 1245
             ||||||                                |||||||||   |||   |||
Db       351 GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGlu 370

Qy      1244 CGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCA------CGC 1191
                        |||    :::    |||      :::            :::     :::
Db       371 GlyGlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAla 390

Qy      1190 CCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCA 1131
             |||||||||   |||   |||:::                               :::|||
Db       391 ProGlyThr---ProGlyProGlnGly-------------------------LeuPro 400

Qy      1130 TCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACA 1071
                   |||      |||                   |||
Db       401 GlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 420

Qy      1070 GCTGGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCC---GGAAAGCCAAGCGGGCCCGGC 1014
                   |||      |||      |||   |||||||||   |||   |||      ||||||
Db       421 ProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly 440

Qy      1013 AT-GGACAGCAG---------TGGGGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTGGC 964
                |||||||||            |||   |||:::       |||      |||   |||
Db       441 GluGlyGlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGly 460

Qy       963 TCGGTGGGG-----------------------------CCCAGCGCTGCCTCCTCAGCC 934
             :::   |||                               |||   |||   |||   ||
Db       461 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla-ProGlyThrPr 480

Qy       933 ACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGCAC 874
             |        |||                   |||        |||      |||   :::|||
Db       480 oGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLe 500

Qy       873 TCCT---------------CCTGGGTGCCCAGGTAGGGGGCCAGGGCACTGGTGTCCCAG 829
              |||                  ||||||   ||||||   |||      |||
Db       500 uProGlySerProGlyAlaProGlyThrProGlyGluGlyGlnGlnHisHisLeuGlyGl 520

Qy       828 TCAATGGCAGGCAGGA------GGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCA 775
                   |||||||||         |||||||||||:::|||
Db       520 yAlaLysGlnAlaGlyAspValGlySerProGlyAlaPro------------------- 533

Qy       774 TAGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGT---CCCGGAAGAGGTCAGAGAGC 721
                            ||||||      |||      ||||||   ||||||
Db       534 -------GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPr 551

Qy       720 AGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCAGCAGCCCCACGCCCAGG 661
              ||||||                   :::||||||           :::||    |||   |||
Db       551 oGlyProGlnGlyLeuProGlySerProGly------AlaProGlyThrProGlyProGl 569

Qy       660 ATGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCG 601
                         :::|||||||||      |||||||||
Db       569 nGlyLeuProGlySerProGlyAla------ProGlyThrProGlyGluGlyGlnGlnHi 587

Qy       600 GCCCTTGGGATGAGAAAGA---------------GGCTCAGCAGGATGCCCAAGGACAGT 556
                ||||||       :::                   ||||||   |||   |||   |||
Db       587 sHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAlaProGlyThrPr 607

Qy       555 G---CCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAG 499
                |||:::        |||                                      ||
Db       607 oGlyProGlnGlyLeuProGly----------------------------------Se 615

Qy       498 CCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCATGGTCATG 439
             |   ||||||     |||||||||
Db       615 rProGlyAlaProGlyThrProGlyPro----------------------------- 624

Qy       438 AACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCC 379
                    ||||||    |||   |||                             ||
Db       625 ----GlnGlyLeuProGlySerProGlyAla--------------------------Pr 634

Qy       378 AAACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGT-- 321
             |  ||||||     |||   |||   |||   |||          |||      |||
Db       634 oGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGl 654

Qy       320 ---------------------------------GCCGCAGCAGGCGGCTCACCCACAGC 295
                                                |||      |||||||||  ||
Db       654 uGlyGlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAl 674
```

```
Qy       1190 CCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCA 1131
              ||||||||    |||    |||:::                              :::|||
Db        391 ProGlyThr---ProGlyProGlnGly-------------------------LeuPro 400

Qy       1130 TCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACA 1071
                 |||      |||                     |||
Db        401 GlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 420

Qy       1070 GCTGGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCC---GGAAAGCCAAGCGGGCCCGGC 1014
                 |||      |||      |||   |||||||||    |||    |||      ||||||
Db        421 ProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly 440

Qy       1013 AT-GGACAGCAG---------TGGGGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTGGC 964
                 |||||||||            |||    |||:::      |||       |||   |||
Db        441 GluGlyGlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGly 460

Qy        963 TCGGTGGGG-----------------------------CCCAGCGCTGCCTCCTCAGCC 934
              :::    |||                            |||   |||  |||      ||
Db        461 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla-ProGlyThrPr 480

Qy        933 ACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGCAC 874
              |               |||            |||      |||      |||   :::|||
Db        480 oGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLe 500

Qy        873 TCCT--------------CCTGGGTGCCCAGGTAGGGGGCCAGGGCACTGGTGTCCCAG 829
                |||                ||||||   ||||||    |||      |||
Db        500 uProGlySerProGlyAlaProGlyThrProGlyGluGlyGlnGlnHisHisLeuGlyGl 520

Qy        828 TCAATGGCAGGCAGGA------GGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCA 775
                    |||||||||       ||||||||||||:::|||
Db        520 yAlaLysGlnAlaGlyAspValGlySerProGlyAlaPro------------------- 533

Qy        774 TAGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGT---CCCGGAAGAGGTCAGAGAGC 721
                          ||||||     |||   ||||||   ||||||
Db        534 -------GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPr 551

Qy        720 AGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGCCCAGG 661
                ||||||                :::||||||      :::|||     |||   |||
Db        551 oGlyProGlnGlyLeuProGlySerProGly------AlaProGlyThrProGlyProGl 569

Qy        660 ATGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCG 601
                          :::|||||||||      |||||||||
Db        569 nGlyLeuProGlySerProGlyAla------ProGlyThrProGlyGluGlyGlnGlnHi 587

Qy        600 GCCCTTGGGATGAGAAAGA---------------GGCTCAGCAGGATGCCCAAGGACAGT 556
                 ||||||     :::                   ||||||   |||    |||    |||
Db        587 sHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAlaProGlyThrPr 607

Qy        555 G---CCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAG 499
                 |||:::      |||                                        ||
Db        607 oGlyProGlnGlyLeuProGly----------------------------------Se 615

Qy        498 CCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCATGGTCATG 439
              |  ||||||        |||||||||
Db        615 rProGlyAlaProGlyThrProGlyPro------------------------------ 624

Qy        438 AACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCC 379
                          ||||||     |||   |||                          ||
Db        625 ----GlnGlyLeuProGlySerProGlyAla--------------------------Pr 634

Qy        378 AAACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGT-- 321
              |  ||||||      |||   |||   |||   |||            |||      |||
Db        634 oGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGl 654

Qy        320 --------------------------------GCCGCAGCAGGCGGCTCACCCACAGC 295
                                              |||        |||||||||  ||
Db        654 uGlyGlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAl 674

Qy        294 CTCTGGACCATAGTGGGCCAGGCGGG---TAGGGCTCAGGGGGCCGTTCAGGCACTCCAG 238
              |  |||      ||||||   |||      ||||||   |||      |||||||||
Db        674 aProGlyThr-ProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProG 694

Qy        237 AACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAAC 178
                          |||         :::|||    |||         :::|||
Db        694 lyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly------------ 709

Qy        177 TGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGCTTCTC 118
                |||:::        ||||||:::                 |||         |||     |
Db        710 --ProGlnGlyLeuProGlySerProGlyAla---ProGlyThrProGlyGluGlyGlnG 728

Qy        117 AGCCCATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTCATCA---- 62
                ||      |||                    ||||||   |||    |||    |||    |||
Db        728 lnHisHisLeuGly-----------GlyAlaLysGlnAlaGlyAspValGlySerProG 744

Qy         61 -----CTCAGATCCTGGCCGA 46
                     ::: ||||||||||||
Db        744 lyAlaMet-AspProGlyArg 750
```

RESULT 17
US-08-397-633A-26
; Sequence 26, Application US/08397633A
; Patent No. 5773577
; GENERAL INFORMATION:
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: PRODUCTS COMPRISING SUBSTRATESCAPABLE
; TITLE OF INVENTION: OF ENZYMATIC CROSS-LINKING
; NUMBER OF SEQUENCES: 105
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/397,633A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A-58848-1/BIR PROP-011-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 762 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-397-633A-26

Alignment Scores:

Pred. No.:	5.68e-22	Length:	762
Score:	391.00	Matches:	259
Percent Similarity:	35.17%	Conservative:	41
Best Local Similarity:	30.36%	Mismatches:	296
Query Match:	6.28%	Indels:	259
DB:	1	Gaps:	51

US-09-759-143-110 (1-3410) x US-08-397-633A-26 (1-762)

```
Qy       2347 GGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCT--- 2291
                |||  ||||||                        ||||||    |||    |||
Db         76 GlySerProGly--------------AlaProGlyThrProGlyGluGlyGlnGlnHis 90

Qy       2290 CAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCA---------GAG 2240
```

```
Db        709 --ProGlnGlyLeuProGlySerProGlyAla---ProGlyThrProGlyGluGlyGlnG 727

Qy        117 AGCCCATGCTCAACACCTGCTGCTGTGGGCACCTCAGTGGGGACACGTCTCATCA---- 62
              ||      |||                  ||||||  |||   |||   |||   |||
Db        727 lnHisHisLeuGly-----------GlyAlaLysGlnAlaGlyAspValGlySerProG 743

Qy         61 -----CTCAGATCCTGGCCGA 46
                   ::: ||||||||||||
Db        743 lyAlaMet-AspProGlyArg 749
```

RESULT 16
US-08-642-255-114
; Sequence 114, Application US/08642255
; Patent No. 5773249
; GENERAL INFORMATION:
;   APPLICANT: CAPPELLO, Joseph
;   APPLICANT: FERRARI, Franco A.
;   TITLE OF INVENTION: High Molecular Weight Collagen-Like
;   TITLE OF INVENTION: Protein Polymers
;   NUMBER OF SEQUENCES: 135
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
;   STREET: 4 Embarcadero Center, Suite 3400
;   CITY: San Francisco
;   STATE: California
;   COUNTRY: USA
;   ZIP: 94111-4187
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/642,255
;   FILING DATE:
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: ROWLAND, Bertram I.
;   REGISTRATION NUMBER: 20,015
;   REFERENCE/DOCKET NUMBER: A55556-3/BIR
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415) 494-8700
;   TELEFAX: (415) 494-8771
;   TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 114:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 762 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-642-255-114

Alignment Scores:

Pred. No.:	5.68e-22	Length:	762
Score:	391.00	Matches:	259
Percent Similarity:	35.17%	Conservative:	41
Best Local Similarity:	30.36%	Mismatches:	296
Query Match:	6.28%	Indels:	259
DB:	1	Gaps:	51

US-09-759-143-110 (1-3410) x US-08-642-255-114 (1-762)

```
Qy       2347 GGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCT--- 2291
               |||  ||||||                |||||      |||    |||
Db         76 GlySerProGly--------------AlaProGlyThrProGlyGluGlyGlnGlnHis 90

Qy       2290 CAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCA---------GAG 2240
                   |||||||||                 :::|||:::|||   |||
Db         91 HisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAlaProGlyThrPro 110

Qy       2239 TCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACC 2180
                  |||              ||||||  |||   |||              |||
Db        111 GlyProGlnGlyLeuProGlySerPro---GlyAlaPro---------GlyThrProGly 126

Qy       2179 CCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCCC 2120
              |||   :::||||||      ||||||||                   ||||||||
Db        127 Pro---GlnGlyLeu---ProGlySerPro-----------GlyAlaProGlyThrPro 140

Qy       2119 CAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGT------GGCAGCAGAGAG 2066
               |||      :::          :::      |||   |||      |||      :::
Db        141 GlyProGlnGlyLeuPro---GlySerProGlyAlaProGlyThrProGlyGluGlyGln 159

Qy       2065 CCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAAC 2006
               |||:::   |||                                    ||||||
Db        160 GlnHisHisLeuGly-----------------------------------GlyAlaLys 167

Qy       2005 AGGAGCGGGGAGCTGGGACCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTT 1946
              :::::|||::::::|||             |||   |||   ||||||
Db        168 GlnAlaGlyAspValGlySerPro--GlyAlaProGlyThrPro--------------- 181

Qy       1945 CTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAAT 1886
                                          |||   ||||||
Db        182 -------------------------GlyProGlnGlyLeuPro--------------- 187

Qy       1885 GGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTG 1826
                        ||||||        ||||||     |||
Db        188 -----GlySerProGlyAlaProGlyThrProGlyProGln------------------ 199

Qy       1825 GACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCACTATCCAG 1766
                                          |||   ||||||:::   |||       ||||
Db        200 -------------------------GlyLeuProGlySerProGlyAla------ProG 209

Qy       1765 GATGGCGAGGTCCAGGCAGATGCCCCGGC-----------CCGGAACCACCCTGGCCTC 1718
              ||     ||||||       ||||||           |||       ||||||
Db        209 lyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro-GlyThrProGlyGlu 228

Qy       1717 GGTGGGCTCACCCACCACCACACGTACGGAGACATCACAGGCA---------GAGGCCCC 1667
              |||             ||||||     |||      ||||||          :::||
Db        229 Gly------GlnGlnHisHisLeuGlyGly-AlaLysGlnAlaGlyAspValGlySerPr 246

Qy       1666 GCAGAGCGCGGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGG 1607
              |   :::   |||     |||   :::   ||||||   :::|||
Db        246 oGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro------------- 261

Qy       1606 GAAGGGAGCTCCAGGCTTAGGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCT 1547
                  ||||||||||||        ||||||   :::   :::   |||   |||
Db        262 ----GlyAlaProGlyThr---ProGlyProGlnGlyLeu-ProGlySerPro------- 276

Qy       1546 AGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGTAGAG 1487
                                              |||   |||   :::||||||   :::|
Db        277 ------------------------------GlyAlaProGlyThrProGlyProGlnG 286

Qy       1486 GGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGC 1427
              ||   |||            ||||||   |||        |||   |||||||||||
Db        286 lyLeuPro-----------GlySerProGlyAla-ProGlyThrProGlyGluGly--- 300

Qy       1426 TGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACCGGCAG---------- 1377
                       ||||||                   ||||||         |||
Db        301 ---GlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAla 319

Qy       1376 --------------CCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGC 1332
                             |||||||||   ||||||   |||
Db        320 ProGlyThrProGlyProGlnGly---LeuProGlySerPro----------------- 332

Qy       1331 CGAATCGCTGCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCA 1272
                       ||||||   |||   |||:::      ||||||
Db        333 ------GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro 350

Qy       1271 GGAACA--------------------------------GCCCCAGGCTGCCCATCCGAA 1245
              ||||||                                ||||||||   |||   |||
Db        351 GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGlu 370

Qy       1244 CGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCA------CGC 1191
                       |||   :::   |||   :::              :::   :::
Db        371 GlyGlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAla 390
```

```
Qy    2119 CAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGT------GGCAGCAGAGAG 2066
Db     140 GlyProGlnGlyLeuPro---GlySerProGlyAlaProGlyThrProGlyGluGlyGln 158

Qy    2065 CCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAAC 2006
Db     159 GlnHisHisLeuGly---------------------------------GlyAlaLys 166

Qy    2005 AGGAGCGGGGAGCTGGGACCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTT 1946
Db     167 GlnAlaGlyAspValGlySerPro--GlyAlaProGlyThrPro--------------- 180

Qy    1945 CTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAAT 1886
Db     181 ------------------------GlyProGlnGlyLeuPro---------------- 186

Qy    1885 GGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTG 1826
Db     187 -----GlySerProGlyAlaProGlyThrProGlyProGln------------------ 198

Qy    1825 GACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCACTATCCAG 1766
Db     199 ------------------------GlyLeuProGlySerProGlyAla------ProG 208

Qy    1765 GATGGCGAGGTCCAGGCAGATGCCCCGGC------------CCGGAACCACCCTGGCCTC 1718
Db     208 lyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro-GlyThrProGlyGlu 227

Qy    1717 GGTGGGCTCACCCACCACCACACGTACGGAGACATCACAGGCA---------GAGGCCCC 1667
Db     228 Gly------GlnGlnHisHisLeuGlyGly-AlaLysGlnAlaGlyAspValGlySerPr 245

Qy    1666 GCAGAGCGCGGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGG 1607
Db     245 oGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro------------- 260

Qy    1606 GAAGGGAGCTCCAGGCTTAGGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCT 1547
Db     261 ----GlyAlaProGlyThr---ProGlyProGlnGlyLeu-ProGlySerPro------- 275

Qy    1546 AGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGTAGAG 1487
Db     276 -------------------------------GlyAlaProGlyThrProGlyProGlnG 285

Qy    1486 GGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGC 1427
Db     285 lyLeuPro-----------GlySerProGlyAla-ProGlyThrProGlyGluGly--- 299

Qy    1426 TGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACCGGCAG---------- 1377
Db     300 ---GlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAla 318

Qy    1376 ---------------CCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGC 1332
Db     319 ProGlyThrProGlyProGlnGly---LeuProGlySerPro----------------- 331

Qy    1331 CGAATCGCTGCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCA 1272
Db     332 ------GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro 349

Qy    1271 GGAACA-------------------------------GCCCCAGGCTGCCCATCCGAA 1245
Db     350 GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGlu 369

Qy    1244 CGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCA------CGC 1191
Db     370 GlyGlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAla 389

Qy    1190 CCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCA 1131
Db     390 ProGlyThr---ProGlyProGlnGly-------------------------LeuPro 399

Qy    1130 TCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACA 1071
Db     400 GlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 419

Qy    1070 GCTGGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCC---GGAAAGCCAAGCGGGCCCGGC 1014
Db     420 ProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly 439

Qy    1013 AT-GGACAGCAG---------TGGGGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTGGC 964
Db     440 GluGlyGlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGly 459

Qy     963 TCGGTGGGG-----------------------------CCCAGCGCTGCCTCCTCAGCC 934
Db     460 AlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla-ProGlyThrPr 479

Qy     933 ACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGCAC 874
Db     479 oGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyProGlnGlyLe 499

Qy     873 TCCT---------------CCTGGGTGCCCAGGTAGGGGGCCAGGGCACTGGTGTCCCAG 829
Db     499 uProGlySerProGlyAlaProGlyThrProGlyGluGlyGlnGlnHisHisLeuGlyGl 519

Qy     828 TCAATGGCAGGCAGGA------GGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCA 775
Db     519 yAlaLysGlnAlaGlyAspValGlySerProGlyAlaPro------------------- 532

Qy     774 TAGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGT---CCCGGAAGAGGTCAGAGAGC 721
Db     533 -------GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrPr 550

Qy     720 AGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGCCCAGG 661
Db     550 oGlyProGlnGlyLeuProGlySerProGly------AlaProGlyThrProGlyProGl 568

Qy     660 ATGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCG 601
Db     568 nGlyLeuProGlySerProGlyAla------ProGlyThrProGlyGluGlyGlnGlnHi 586

Qy     600 GCCCTTGGGATGAGAAAGA--------------GGCTCAGCAGGATGCCCAAGGACAGT 556
Db     586 sHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAlaProGlyThrPr 606

Qy     555 G---CCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAG 499
Db     606 oGlyProGlnGlyLeuProGly-----------------------------------Se 614

Qy     498 CCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCATGGTCATG 439
Db     614 rProGlyAlaProGlyThrProGlyPro------------------------------ 623

Qy     438 AACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCC 379
Db     624 ----GlnGlyLeuProGlySerProGlyAla--------------------------Pr 633

Qy     378 AAACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGT-- 321
Db     633 oGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGl 653

Qy     320 ---------------------------------GCCGCAGCAGGCGGCTCACCCACAGC 295
Db     653 uGlyGlnGlnHisHisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAl 673

Qy     294 CTCTGGACCATAGTGGGCCAGGCGGG---TAGGGCTCAGGGGGCCGTTCAGGCACTCCAG 238
Db     673 aProGlyThr-ProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProG 693

Qy     237 AACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAAC 178
Db     693 lyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly------------ 708

Qy     177 TGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGCTTCTC 118
```

```
Qy        782 TGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGT---CCCGGAAGAGGT 729
                                |||||| |||  |||||| ||||||
Db        534 ---------------GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro 548

Qy        728 CAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCA 669
                         ||||||            :::||||||      :::|||    |||
Db        549 GlyThrProGlyProGlnGlyLeuProGlySerProGly------AlaProGlyThrPro 566

Qy        668 CGCCCAGGATGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTA 609
                 |||            :::|||||||||     |||||||||
Db        567 GlyProGlnGlyLeuProGlySerProGlyAla------ProGlyThrProGlyGluGly 584

Qy        608 GCCAGCCGGCCCTTGGGATGAGAAAGA---------------GGCTCAGCAGGATGCCCA 564
                          ||||||  |||           ||||||  |||  |||
Db        585 GlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaPro 604

Qy        563 AGGACAGTG---CCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCAC 507
                 |||    |||:::       |||
Db        605 GlyThrProGlyProGlnGlyLeuProGly------------------------------ 614

Qy        506 TGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCA 447
                    |||  ||||||       |||||||||
Db        615 ------SerProGlyAlaProGlyThrProGlyPro----------------------- 624

Qy        446 TGGTCATGAACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGC 387
                                ||||||    |||  |||
Db        625 ------------GlnGlyLeuProGlySerProGlyAla--------------------- 633

Qy        386 CTGCGGCCAAACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTT 327
                    |||  ||||||   |||  |||  |||  |||          |||
Db        634 ------ProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr 651

Qy        326 TCCGGT----------------------------------GCCGCAGCAGGCGGCTCA 303
                 |||                                   |||       ||||||
Db        652 ProGlyGluGlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySer 671

Qy        302 CCCACAGCCTCTGGACCATAGTGGGCCAGGCGGG---TAGGGCTCAGGGGGCCGTTCAGG 246
              |||  |||  |||    ||||||  |||   ||||||  |||    ||
Db        672 ProGlyAlaProGlyThr-ProGlyProGlnGlyLeuProGlySerProGlyAlaProGl 691

Qy        245 CACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTG 186
              |||||||        |||    :::|||  |||     :::|||
Db        691 yThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly----- 709

Qy        185 CCGCCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCC 126
                          |||:::  ||||||:::          |||         |||
Db        710 ---------ProGlnGlyLeuProGlySerProGlyAla---ProGlyThrProGlyGl 725

Qy        125 AGCTTCTCAGCCCATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTC 66
                   |||    |||               ||||||  |||  |||  |||
Db        725 uGlyGlnGlnHisHisLeuGly-----------GlyAlaArgGlnAlaGlyAspValGl 741

Qy         65 ATCA---------CTCAGATCCTGGCCGA 46
               |||         ::: |||||||||||
Db        741 ySerProGlyAlaMet-AspProGlyArg 750
```

RESULT 15
US-08-707-237A-84
; Sequence 84, Application US/08707237A
; Patent No. 5830713
; GENERAL INFORMATION:
;   APPLICANT: Ferrari, Franco A.
;   APPLICANT: Capello, Joseph
;   APPLICANT: Crissman, John W.
;   APPLICANT: Dorman, Mary A.
;   TITLE OF INVENTION: METHODS FOR PREPARING SYNTHETIC
;   TITLE OF INVENTION: REPETITIVE DNA
;   NUMBER OF SEQUENCES: 108
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
;   STREET: Four Embarcadero Center, Suite 3400
;   CITY: San Francisco
;   STATE: California
;   COUNTRY: United States
;   ZIP: 94111-4187
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/707,237A
;   FILING DATE: 03-SEP-1996
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/175,155
;   FILING DATE: 29-DEC-1993
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/053,049
;   FILING DATE: 22-APR-1993
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/609,716
;   FILING DATE: 06-NOV-1990
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/269,429
;   FILING DATE: 09-NOV-1988
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/114,618
;   FILING DATE: 29-OCT-1987
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 06/927,258
;   FILING DATE: 04-NOV-1986
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Trecartin, Richard F.
;   REGISTRATION NUMBER: 31,801
;   REFERENCE/DOCKET NUMBER: A-55186-10/WHD
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415) 781-1989
;   TELEFAX: (415) 398-3249
;   TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 84:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 761 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: unknown
;   TOPOLOGY: unknown
;   MOLECULE TYPE: protein
US-08-707-237A-84

Alignment Scores:

Pred. No.:	3.99e-22	Length:	761
Score:	393.00	Matches:	259
Percent Similarity:	35.17%	Conservative:	41
Best Local Similarity:	30.36%	Mismatches:	296
Query Match:	6.31%	Indels:	259
DB:	2	Gaps:	51

US-09-759-143-110 (1-3410) x US-08-707-237A-84 (1-761)

```
Qy       2347 GGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCT--- 2291
                 |||  ||||||               ||||||    |||   |||
Db         75 GlySerProGly---------------AlaProGlyThrProGlyGluGlyGlnGlnHis 89

Qy       2290 CAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCA---------GAG 2240
                 |||||||||                     :::|||:::|||   |||
Db         90 HisLeuGlyGlyAlaLysGlnAlaGlyAspValGlySerProGlyAlaProGlyThrPro 109

Qy       2239 TCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACC 2180
                 |||                  ||||||  |||   |||          |||
Db        110 GlyProGlnGlyLeuProGlySerPro---GlyAlaPro---------GlyThrProGly 125

Qy       2179 CCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCCC 2120
                 |||   :::||||||      |||||||||                 |||||||||
Db        126 Pro---GlnGlyLeu---ProGlySerPro------------GlyAlaProGlyThrPro 139
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/397,633A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rowland, Bertram I
;      REGISTRATION NUMBER:  20,015
;      REFERENCE/DOCKET NUMBER:  A-58848-1/BIR PROP-011-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 781-1989
;      TELEFAX:  (415) 398-3249
;      TELEX:  910 277299
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  762 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-397-633A-31

Alignment Scores:
Pred. No.:               2.8e-22      Length:         762
Score:                   395.00       Matches:        257
Percent Similarity:      34.50%       Conservative:   38
Best Local Similarity:   30.06%       Mismatches:     298
Query Match:             6.35%        Indels:         263
DB:                      1            Gaps:           52

US-09-759-143-110 (1-3410) x US-08-397-633A-31 (1-762)

Qy     2347 GGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCT--- 2291
            |||  ||||||                      ||||||    |||     |||
Db       76 GlySerProGly---------------AlaProGlyThrProGlyGluGlyGlnGlnHis 90

Qy     2290 CAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCA---------GAG 2240
               |||||||||                    :::|||:::|||   |||
Db       91 HisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaProGlyThrPro 110

Qy     2239 TCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCAGACTGAAACC 2180
               |||              ||||||  |||  |||          |||
Db      111 GlyProGlnGlyLeuProGlySerPro---GlyAlaPro---------GlyThrProGly 126

Qy     2179 CCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCCC 2120
            |||  :::||||||    |||||||||                   |||||||||
Db      127 Pro---GlnGlyLeu---ProGlySerPro-----------GlyAlaProGlyThrPro 140

Qy     2119 CAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGT------GGCAGCAGAGAG 2066
               |||     :::         :::     |||  |||    |||       :::
Db      141 GlyProGlnGlyLeuPro---GlySerProGlyAlaProGlyThrProGlyGluGlyGln 159

Qy     2065 CCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAAC 2006
               |||:::  |||                                   ||||||
Db      160 GlnHisHisLeuGly----------------------------------GlyAlaArg 167

Qy     2005 AGGAGCGGGGAGCTGGGACCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTT 1946
            ::::::|||::::::|||         |||   |||   ||||||
Db      168 GlnAlaGlyAspValGlySerPro--GlyAlaProGlyThrPro--------------- 181

Qy     1945 CTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAAT 1886
                                              |||   ||||||
Db      182 -------------------------GlyProGlnGlyLeuPro---------------- 187

Qy     1885 GGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTG 1826
                     ||||||         ||||||     |||
Db      188 -----GlySerProGlyAlaProGlyThrProGlyProGln------------------ 199

Qy     1825 GACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCACTATCCAG 1766
                                         |||   ||||||:::   |||      ||||
Db      200 -------------------------GlyLeuProGlySerProGlyAla------ProG 209

Qy     1765 GATGGCGAGGTCCAGGCAGATGCCCCGGC------------CCGGAACCACCCTGGCCTC 1718
            ||      ||||||          ||||||           |||        ||||||

Db      209 lyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro-GlyThrProGlyGlu 228

Qy     1717 GGTGGGCTCACCCACCACCACACGTACGGAGACATCACAGGCA---------GAGGCCCC 1667
            |||         ||||||     |||        ||||||          :::||
Db      229 Gly------GlnGlnHisHisLeuGlyGly-AlaArgGlnAlaGlyAspValGlySerPr 246

Qy     1666 GCAGAGCGCGGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGG 1607
            |   :::   |||      |||     :::    ||||||    :::|||
Db      246 oGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro------------- 261

Qy     1606 GAAGGGAGCTCCAGGCTTAGGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCT 1547
                |||||||||||        ||||||   :::     :::    |||   |||
Db      262 ----GlyAlaProGlyThr---ProGlyProGlnGlyLeu-ProGlySerPro------- 276

Qy     1546 AGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGTAGAG 1487
                                             |||   |||   :::||||||   :::|
Db      277 ------------------------------GlyAlaProGlyThrProGlyProGlnG 286

Qy     1486 GGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGC 1427
            ||   |||               ||||||  |||     |||   |||||||||||
Db      286 lyLeuPro-----------GlySerProGlyAla-ProGlyThrProGlyGluGly--- 300

Qy     1426 TGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACCGGCAG---------- 1377
                      ||||||              ||||||          |||
Db      301 ---GlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAla 319

Qy     1376 --------------CCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGC 1332
                          |||||||||   ||||||   |||
Db      320 ProGlyThrProGlyProGlnGly---LeuProGlySerPro----------------- 332

Qy     1331 CGAATCGCTGCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCA 1272
                      ||||||    |||   |||:::       ||||||
Db      333 ------GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro 350

Qy     1271 GGAACA------------------------------GCCCCAGGCTGCCCATCCGAA 1245
            ||||||                                  |||||||||   |||   |||
Db      351 GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGlu 370

Qy     1244 CGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCA------CGC 1191
                 |||   :::   |||   |||                  :::    :::
Db      371 GlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAla 390

Qy     1190 CCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCA 1131
            |||||||||   |||   |||:::                              :::|||
Db      391 ProGlyThr---ProGlyProGlnGly----------------------------LeuPro 400

Qy     1130 TCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACA 1071
               |||      |||                        |||
Db      401 GlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 420

Qy     1070 GCTGGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCC---GGAAAGCCAAGCGGGCCCGGC 1014
               |||      |||      |||   |||||||||   |||   |||      ||||||
Db      421 ProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly 440

Qy     1013 -------------------------------------------ATGGACAGCAGTGGG 999
                                                        :::    |||   |||
Db      441 GluGlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGly 460

Qy      998 GCGACA---AGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCT 942
            |||         ||||||     |||                     |||    |||
Db      461 AlaProGlyThrProGlyProGlnGlyLeu---------ProGlySerProGlyAlaPro 477

Qy      941 CCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAA 882
                   |||            |||            |||      |||      |||
Db      478 GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro 497

Qy      881 AGAGGCACTCCT---------------CCTGGGTGCCCAGGTAGGGGGCCAGGGCACTGG 837
            :::|||   |||               ||||||   ||||||   |||      |||
Db      498 GlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGluGlyGlnGlnHisHis 517

Qy      836 TGTCCCAGTCAATGGCAGGCAGGA------GGTAGCCCAGGCAGCCCCCAAGACTGATCA 783
                           |||||||||         ||||||||||||:::|||
Db      518 LeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaPro----------- 533
```

```
Qy      1666 GCAGAGCGCGGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGG 1607
             |   :::   |||     |||   :::   ||||||    :::|||
Db       246 oGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerPro-------------- 261

Qy      1606 GAAGGGAGCTCCAGGCTTAGGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCT 1547
                 |||||||||||      ||||||    :::   :::    |||   |||
Db       262 ----GlyAlaProGlyThr---ProGlyProGlnGlyLeu-ProGlySerPro------- 276

Qy      1546 AGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGTAGAG 1487
                                            |||   |||   :::||||||   :::|
Db       277 -------------------------------GlyAlaProGlyThrProGlyProGlnG 286

Qy      1486 GGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGC 1427
             ||   |||                ||||||   |||       |||   |||||||||||
Db       286 lyLeuPro-----------GlySerProGlyAla-ProGlyThrProGlyGluGly--- 300

Qy      1426 TGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACCGGCAG---------- 1377
                       ||||||                ||||||       |||
Db       301 ---GlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAla 319

Qy      1376 --------------CCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGC 1332
                           |||||||||   ||||||     |||
Db       320 ProGlyThrProGlyProGlnGly---LeuProGlySerPro----------------- 332

Qy      1331 CGAATCGCTGCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCA 1272
                       ||||||     |||    |||:::      ||||||
Db       333 ------GlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro 350

Qy      1271 GGAACA---------------------------------GCCCCAGGCTGCCCATCCGAA 1245
             ||||||                                 |||||||||   |||   |||
Db       351 GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGlu 370

Qy      1244 CGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCA------CGC 1191
                      |||   :::   |||   |||             :::    :::
Db       371 GlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAla 390

Qy      1190 CCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCA 1131
             |||||||||   |||   |||:::                           :::|||
Db       391 ProGlyThr---ProGlyProGlnGly--------------------------LeuPro 400

Qy      1130 TCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACA 1071
                |||      |||                   |||
Db       401 GlySerProGlyAlaProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAla 420

Qy      1070 GCTGGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCC---GGAAAGCCAAGCGGCCCCGGC 1014
                |||      |||      |||   |||||||||   |||   |||      ||||||
Db       421 ProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly 440

Qy      1013 ------------------------------------------ATGGACAGCAGTGGG 999
                                                       :::   |||   |||
Db       441 GluGlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGly 460

Qy       998 GCGACA---AGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCT 942
             |||            ||||||      |||                    |||   |||
Db       461 AlaProGlyThrProGlyProGlnGlyLeu---------ProGlySerProGlyAlaPro 477

Qy       941 CCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAA 882
                   |||         |||            |||      |||      |||
Db       478 GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyPro 497

Qy       881 AGAGGCACTCCT--------------CCTGGGTGCCCAGGTAGGGGGCCAGGGCACTGG 837
             :::|||   |||              ||||||   ||||||      |||      |||
Db       498 GlnGlyLeuProGlySerProGlyAlaProGlyThrProGlyGluGlyGlnGlnHisHis 517

Qy       836 TGTCCCAGTCAATGGCAGGCAGGA------GGTAGCCCAGGCAGCCCCCAAGACTGATCA 783
                                ||||||||||       ||||||||||||:::|||
Db       518 LeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaPro----------- 533

Qy       782 TGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGT---CCCGGAAGAGGT 729
                                        ||||||      |||   ||||||    ||||||
Db       534 --------------GlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro 548

Qy       728 CAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCA 669
                       ||||||                 :::||||||      :::|||    |||
Db       549 GlyThrProGlyProGlnGlyLeuProGlySerProGly------AlaProGlyThrPro 566

Qy       668 CGCCCAGGATGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTA 609
                |||                :::|||||||||       |||||||||
Db       567 GlyProGlnGlyLeuProGlySerProGlyAla------ProGlyThrProGlyGluGly 584

Qy       608 GCCAGCCGGCCCTTGGGATGAGAAAGA--------------GGCTCAGCAGGATGCCCA 564
                              ||||||      |||                ||||||   |||   |||
Db       585 GlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaPro 604

Qy       563 AGGACAGTG---CCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCAC 507
                |||      |||:::        |||
Db       605 GlyThrProGlyProGlnGlyLeuProGly----------------------------- 614

Qy       506 TGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCA 447
                   |||   ||||||          |||||||||
Db       615 ------SerProGlyAlaProGlyThrProGlyPro----------------------- 624

Qy       446 TGGTCATGAACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGC 387
                              ||||||    |||   |||
Db       625 -----------GlnGlyLeuProGlySerProGlyAla------------------- 633

Qy       386 CTGCGGCCAAACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTT 327
                   |||   ||||||      |||   |||   |||   |||         |||
Db       634 ------ProGlyThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThr 651

Qy       326 TCCGGT------------------------------------GCCGCAGCAGGCGGCTCA 303
                |||                                      |||         ||||||
Db       652 ProGlyGluGlyGlnGlnHisHisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySer 671

Qy       302 CCCACAGCCTCTGGACCATAGTGGGCCAGGCGGG---TAGGGCTCAGGGGGCCGTTCAGG 246
               |||   |||   |||         ||||||   |||         ||||||   |||   ||
Db       672 ProGlyAlaProGlyThr-ProGlyProGlnGlyLeuProGlySerProGlyAlaProGl 691

Qy       245 CACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTG 186
             ||||||         |||         :::|||   |||         :::|||
Db       691 yThrProGlyProGlnGlyLeuProGlySerProGlyAlaProGlyThrProGly----- 709

Qy       185 CCGCCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCC 126
                               |||:::      ||||||:::                |||         |||
Db       710 ---------ProGlnGlyLeuProGlySerProGlyAla---ProGlyThrProGlyGl 725

Qy       125 AGCTTCTCAGCCCATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTC 66
                   |||      |||                  ||||||   |||   |||   |||
Db       725 uGlyGlnGlnHisHisLeuGly-----------GlyAlaArgGlnAlaGlyAspValGl 741

Qy        65 ATCA---------CTCAGATCCTGGCCGA 46
                |||         :::  ||||||||||||
Db       741 ySerProGlyAlaMet-AspProGlyArg 750
```

RESULT 14
US-08-397-633A-31
; Sequence 31, Application US/08397633A
; Patent No. 5773577
; GENERAL INFORMATION:
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: PRODUCTS COMPRISING SUBSTRATESCAPABLE
; TITLE OF INVENTION: OF ENZYMATIC CROSS-LINKING
; NUMBER OF SEQUENCES: 105
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```
; APPLICANT:  Wang, Aijun
; APPLICANT:  Skeiky, Yasir A.W.
; APPLICANT:  Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C16
; CURRENT APPLICATION NUMBER: US/09/605,785
; CURRENT FILING DATE:  2000-06-27
; NUMBER OF SEQ ID NOS: 835
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 706
;   LENGTH: 123
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-605-785-706

Alignment Scores:
Pred. No.:                6.16e-25     Length:          123
Score:                    426.00       Matches:         87
Percent Similarity:       97.83%       Conservative:    3
Best Local Similarity:    94.57%       Mismatches:      2
Query Match:              6.64%        Indels:          0
DB:                       4            Gaps:            0

US-09-759-143-110 (1-3410) x US-09-605-785-706 (1-123)

Qy       1250 ATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATG 1309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeuValMet 20

Qy       1310 GACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTC 1369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAlaAlaPhe 40

Qy       1370 CCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCC 1429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAlaSerAla 60

Qy       1430 GCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTC 1489
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAlaSerLeu 80

Qy       1490 TACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATAC 1525
              ||||||||||||||||||||   :::   ::::::
Db         81 TyrHisArgGluLysGlnValLeuIleGlyGlnTrp 92

RESULT 13
US-08-642-255-120
; Sequence 120, Application US/08642255
; Patent No. 5773249
;  GENERAL INFORMATION:
;    APPLICANT:  CAPPELLO, Joseph
;    APPLICANT:  FERRARI, Franco A.
;    TITLE OF INVENTION:  High Molecular Weight Collagen-Like
;    TITLE OF INVENTION:  Protein Polymers
;    NUMBER OF SEQUENCES:  135
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/642,255
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROWLAND, Bertram I.
;      REGISTRATION NUMBER:  20,015
;      REFERENCE/DOCKET NUMBER:  A55556-3/BIR
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 494-8700
;      TELEFAX:  (415) 494-8771
;      TELEX:  910 277299 FHT UR
;  INFORMATION FOR SEQ ID NO:  120:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  762 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-642-255-120

Alignment Scores:
Pred. No.:                2.8e-22      Length:          762
Score:                    395.00       Matches:         257
Percent Similarity:       34.50%       Conservative:    38
Best Local Similarity:    30.06%       Mismatches:      298
Query Match:              6.35%        Indels:          263
DB:                       1            Gaps:            52

US-09-759-143-110 (1-3410) x US-08-642-255-120 (1-762)

Qy       2347 GGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCT--- 2291
               |||  ||||||                   ||||||      |||      |||
Db         76 GlySerProGly---------------AlaProGlyThrProGlyGluGlyGlnGlnHis 90

Qy       2290 CAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCA---------GAG 2240
                 ||||||||               :::|||:::|||   |||
Db         91 HisLeuGlyGlyAlaArgGlnAlaGlyAspValGlySerProGlyAlaProGlyThrPro 110

Qy       2239 TCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACC 2180
                 |||              ||||||   |||   |||             |||
Db        111 GlyProGlnGlyLeuProGlySerPro---GlyAlaPro---------GlyThrProGly 126

Qy       2179 CCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCCC 2120
               |||   :::||||||      |||||||||                  |||||||||
Db        127 Pro---GlnGlyLeu---ProGlySerPro------------GlyAlaProGlyThrPro 140

Qy       2119 CAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGT------GGCAGCAGAGAG 2066
               |||      :::         :::   |||   |||      |||      :::
Db        141 GlyProGlnGlyLeuPro---GlySerProGlyAlaProGlyThrProGlyGluGlyGln 159

Qy       2065 CCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAAC 2006
                 |||:::   |||                                     ||||||
Db        160 GlnHisHisLeuGly---------------------------------GlyAlaArg 167

Qy       2005 AGGAGCGGGGAGCTGGGACCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTT 1946
              ::::::|||::::::|||         |||   |||   ||||||
Db        168 GlnAlaGlyAspValGlySerPro--GlyAlaProGlyThrPro--------------- 181

Qy       1945 CTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAAT 1886
                                                |||  ||||||
Db        182 -------------------------GlyProGlnGlyLeuPro---------------- 187

Qy       1885 GGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTG 1826
                          ||||||      ||||||      |||
Db        188 -----GlySerProGlyAlaProGlyThrProGlyProGln------------------ 199

Qy       1825 GACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCACTATCCAG 1766
                                          |||   ||||||:::   |||      ||||
Db        200 -------------------------GlyLeuProGlySerProGlyAla------ProG 209

Qy       1765 GATGGCGAGGTCCAGGCAGATGCCCCGGC------------CCGGAACCACCCTGGCCTC 1718
               ||      ||||||      ||||||         |||      ||||||
Db        209 lyThrProGlyProGlnGlyLeuProGlySerProGlyAlaPro-GlyThrProGlyGlu 228

Qy       1717 GGTGGGCTCACCCACCACCACACGTACGGAGACATCACAGGCA---------GAGGCCCC 1667
               |||         ||||||   |||   ||||||                   :::||
Db        229 Gly------GlnGlnHisHisLeuGlyGly-AlaArgGlnAlaGlyAspValGlySerPr 246
```

```
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6252047e
US-09-525-397-36

Alignment Scores:
Pred. No.:               7e-91        Length:        255
Score:                   1287.00      Matches:       255
Percent Similarity:      100.00%      Conservative:  0
Best Local Similarity:   100.00%      Mismatches:    0
Query Match:             20.05%       Indels:        0
DB:                      4            Gaps:          0

US-09-759-143-110 (1-3410) x US-09-525-397-36 (1-255)

Qy       1178 GGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGAT 1237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GlyLeuTyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAsp 20

Qy       1238 GAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTC 1297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GluGlyValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPhe 40

Qy       1298 TCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGT 1357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 SerLeuValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSer 60

Qy       1358 GTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTG 1417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ValAlaAlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValVal 80

Qy       1418 ACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACA 1477
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ThrAlaSerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThr 100

Qy       1478 CTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACT 1537
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 LeuAlaSerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThr 120

Qy       1538 GGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGA 1597
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GlyGlyAlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGly 140

Qy       1598 GCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCC 1657
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 AlaProPheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProPro 160

Qy       1658 GCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACC 1717
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AlaLeuCysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThr 180

Qy       1718 GAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCC 1777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GluAlaArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAla 200

Qy       1778 TTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAG 1837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 PheLeuLeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGln 220

Qy       1838 TCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCT 1897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 SerValThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAla 240

Qy       1898 ACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||||||||
Db        241 ThrGlnValValPheAspLysSerAspLeuAlaLysTyrSerAla 255

RESULT 11
US-09-439-313-571
; Sequence 571, Application US/09439313 ;
; Patent No. 6329505
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan Louise
;  APPLICANT:  Jiang Yuqui
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Kalos, Michael
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Retter, Mark
;  APPLICANT:  Solk, John
;  APPLICANT:  Day, Craig
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C9
;  CURRENT APPLICATION NUMBER: US/09/439,313
;  CURRENT FILING DATE:  1999-11-12
;  NUMBER OF SEQ ID NOS: 575
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 571
;   LENGTH: 84
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-439-313-571

Alignment Scores:
Pred. No.:               5.45e-27     Length:        84
Score:                   452.00       Matches:       84
Percent Similarity:      100.00%      Conservative:  0
Best Local Similarity:   100.00%      Mismatches:    0
Query Match:             7.04%        Indels:        0
DB:                      4            Gaps:          0

US-09-759-143-110 (1-3410) x US-09-439-313-571 (1-84)

Qy       1493 CACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGT 1552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 HisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGlyAlaSerSer 20

Qy       1553 GAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAAT 1612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaProPheProAsn 40

Qy       1613 GGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCC 1672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 GlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeuCysGlyAla 60

Qy       1673 TCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTT 1732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SerAlaCysAspValSerValArgValValValGlyGluProThrGluAlaArgValVal 80

Qy       1733 CCGGGCCGGGGC 1744
              ||||||||||||
Db         81 ProGlyArgGly 84

RESULT 12
US-09-605-785-706
; Sequence 706, Application US/09605785
; Patent No. 6321716
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
```

```
;      STREET:  100 Abbott Park Road
;      CITY:  Abbott Park
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/071,710
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/850,713
;      FILING DATE:  02-MAY-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Becker, Cheryl L.
;      REGISTRATION NUMBER:  35,441
;      REFERENCE/DOCKET NUMBER:  6083.US.P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  847/935-1729
;      TELEFAX:  847/938-2623
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  255 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6130043e
US-09-071-710-36

Alignment Scores:
Pred. No.:              7e-91        Length:        255
Score:                  1287.00      Matches:       255
Percent Similarity:     100.00%      Conservative:  0
Best Local Similarity:  100.00%      Mismatches:    0
Query Match:            20.05%       Indels:        0
DB:                     4            Gaps:          0

US-09-759-143-110 (1-3410) x US-09-071-710-36 (1-255)

Qy       1178 GGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGAT 1237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GlyLeuTyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAsp 20

Qy       1238 GAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTC 1297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GluGlyValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPhe 40

Qy       1298 TCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGT 1357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 SerLeuValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSer 60

Qy       1358 GTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTG 1417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ValAlaAlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValVal 80

Qy       1418 ACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACA 1477
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ThrAlaSerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThr 100

Qy       1478 CTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACT 1537
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 LeuAlaSerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThr 120

Qy       1538 GGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGA 1597
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GlyGlyAlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGly 140

Qy       1598 GCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCC 1657
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 AlaProPheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProPro 160

Qy       1658 GCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACC 1717
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AlaLeuCysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThr 180

Qy       1718 GAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCC 1777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GluAlaArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAla 200

Qy       1778 TTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAG 1837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 PheLeuLeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGln 220

Qy       1838 TCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCT 1897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 SerValThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAla 240

Qy       1898 ACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||||||||
Db        241 ThrGlnValValPheAspLysSerAspLeuAlaLysTyrSerAla 255

RESULT 10
US-09-525-397-36
; Sequence 36, Application US/09525397
; Patent No. 6252047
;  GENERAL INFORMATION:
;    APPLICANT:  BILLING-MEDEL, PATRICIA
;    APPLICANT:  COHEN, MAURICE
;    APPLICANT:  COLPITTS, TRACEY L.
;    APPLICANT:  FRIEDMAN, PAULA N.
;    APPLICANT:  GORDON, JULIAN
;    APPLICANT:  GRANADOS, EDWARD N.
;    APPLICANT:  HODGES, STEVEN C.
;    APPLICANT:  KLASS, MICHAEL R.
;    APPLICANT:  KRATOCHVIL, JON D.
;    APPLICANT:  ROBERTS-RAPP, LISA
;    APPLICANT:  RUSSELL, JOHN C.
;    APPLICANT:  STROUPE, STEPHEN D.
;    TITLE OF INVENTION:  REAGENTS AND METHODS USEFUL
;    TITLE OF INVENTION:  FOR DETECTING DISEASES OF THE PROSTATE
;    NUMBER OF SEQUENCES:  41
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  100 Abbott Park Road
;      CITY:  Abbott Park
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/525,397
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/071,710
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Becker, Cheryl L.
;      REGISTRATION NUMBER:  35,441
;      REFERENCE/DOCKET NUMBER:  6083.US.P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  847/935-1729
;      TELEFAX:  847/938-2623
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  255 amino acids
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||
Db        541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 8
US-09-605-785-708
; Sequence 708, Application US/09605785
; Patent No. 6321716
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C16
;  CURRENT APPLICATION NUMBER: US/09/605,785
;  CURRENT FILING DATE:  2000-06-27
;  NUMBER OF SEQ ID NOS: 835
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 708
;   LENGTH: 371
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-605-785-708

Alignment Scores:
Pred. No.:              4.58e-101     Length:        371
Score:                  1420.50       Matches:       275
Percent Similarity:     88.58%        Conservative:  12
Best Local Similarity:  84.88%        Mismatches:    23
Query Match:            22.13%        Indels:        14
DB:                     4             Gaps:          1

US-09-759-143-110 (1-3410) x US-09-605-785-708 (1-371)

Qy.       347 AACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCT 406
              :::|||:::            |||   :::   |||:::      :::|||:::
Db          9 SerLeuValProLeuProLeuAlaLeuTyrLeuSerGlnProLeuThrHisThrThrSer 28

Qy        407 CTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCA 466
              ||||||                                          ||||||||||||
Db         29 LeuLeu------------------------------------------AlaGlyIleGlyPro 35

Qy        467 GTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGC 526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         36 ValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGlyArg 55

Qy        527 TATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTT 586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         56 TyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeuPhe 75

Qy        587 CTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAG 646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 LeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeuGlu 95

Qy        647 CTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACT 706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 LeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPheThr 115

Qy        707 CCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTAC 766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        116 ProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAlaTyr 135

Qy        767 TCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATT 826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 SerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAlaIle 155

Qy        827 GACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGC 886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 AspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPheGly 175

Qy        887 CTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCA 946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 LeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGluAla 195

Qy        947 GCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGC 1006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 AlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHisCys 215

Qy       1007 TGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCAC 1066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 CysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeuHis 235

Qy       1067 CAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGC 1126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 GlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCysSer 255

Qy       1127 TGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTAC 1186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 TrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeuTyr 275

Qy       1187 CAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTT 1246
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 GlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGlyLys 295

Qy       1247 CGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTC 1306
                 ::: ||||||          |||   :::       |||   ::::::|||   ::
Db        296 AlaLeu-AlaAlaSerArgGlyTrpCysGlySerArgProProGluThrThrLeuGlyAl 315

Qy       1307 ATGGACCGGC 1316
               :   :::|||
Db        315 aValSerGly 318

RESULT 9
US-09-071-710-36
; Sequence 36, Application US/09071710
; Patent No. 6130043
;  GENERAL INFORMATION:
;    APPLICANT:  BILLING-MEDEL, PATRICIA
;    APPLICANT:  COHEN, MAURICE
;    APPLICANT:  COLPITTS, TRACEY L.
;    APPLICANT:  FRIEDMAN, PAULA N.
;    APPLICANT:  GORDON, JULIAN
;    APPLICANT:  GRANADOS, EDWARD N.
;    APPLICANT:  HODGES, STEVEN C.
;    APPLICANT:  KLASS, MICHAEL R.
;    APPLICANT:  KRATOCHVIL, JON D.
;    APPLICANT:  ROBERTS-RAPP, LISA
;    APPLICANT:  RUSSELL, JOHN C.
;    APPLICANT:  STROUPE, STEPHEN D.
;    TITLE OF INVENTION:  REAGENTS AND METHODS USEFUL
;    TITLE OF INVENTION:  FOR DETECTING DISEASES OF THE PROSTATE
;    NUMBER OF SEQUENCES:  41
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
```

```
Db        481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||
Db        541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 7
US-09-232-149A-113
; Sequence 113, Application US/09232149A
; Patent No. 6465611
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer Lynn
;  TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE
;  TITLE OF INVENTION:  CANCER AND METHODS FOR THEIR USE
;  FILE REFERENCE: 210121.427C6
;  CURRENT APPLICATION NUMBER: US/09/232,149A
;  CURRENT FILING DATE:  1999-01-15
;  NUMBER OF SEQ ID NOS: 338
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 113
;   LENGTH: 553
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-232-149A-113

Alignment Scores:
Pred. No.:              1.61e-211      Length:         553
Score:                  2861.00        Matches:        553
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            44.58%         Indels:         0
DB:                     4              Gaps:           0

US-09-759-143-110 (1-3410) x US-09-232-149A-113 (1-553)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
```

```
Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||
Db        541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 6
US-09-602-877A-101
; Sequence 101, Application US/09602877A
; Patent No. 6432707
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE:  210121.446C5
;  CURRENT APPLICATION NUMBER: US/09/602,877A
;  CURRENT FILING DATE:  2000-06-22
;  NUMBER OF SEQ ID NOS:  107
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 101
;   LENGTH: 553
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-602-877A-101

Alignment Scores:
Pred. No.:              1.61e-211      Length:         553
Score:                  2861.00        Matches:        553
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            44.58%         Indels:         0
DB:                     4              Gaps:           0

US-09-759-143-110 (1-3410) x US-09-602-877A-101 (1-553)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy      1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy      1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy      1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy      1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
             |||||||||||||||||||||||||||||||||||||||
Db       541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 5
US-09-352-616A-113
; Sequence 113, Application US/09352616A
; Patent No. 6395278
; GENERAL INFORMATION:
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Harlocker, Susan Louise
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Mitcham, Jennifer Lynn
;  TITLE OF INVENTION:  COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
;  TITLE OF INVENTION:  OF PROSTATE CANCER AND METHODS FOR THEIR USE
;  FILE REFERENCE:  210121.427C8
;  CURRENT APPLICATION NUMBER: US/09/352,616A
;  CURRENT FILING DATE:  1999-07-13
;  NUMBER OF SEQ ID NOS:  472
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 113
;   LENGTH: 553
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-352-616A-113

Alignment Scores:
Pred. No.:              1.61e-211     Length:          553
Score:                  2861.00       Matches:         553
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            44.58%        Indels:          0
DB:                     4             Gaps:            0

US-09-759-143-110 (1-3410) x US-09-352-616A-113 (1-553)

Qy       284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy       344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy       404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy       464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy       524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy       584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy       644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy       704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy       764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy       824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy       884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy       944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy      1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy      1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy      1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy      1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy      1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy      1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy      1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy      1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy      1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy      1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy      1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy      1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480
```

```
Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||
Db        541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553
```

RESULT 4
US-09-439-313-113
; Sequence 113, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqui
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439,313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 113
;   LENGTH: 553
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-439-313-113

```
Alignment Scores:
Pred. No.:              1.61e-211     Length:        553
Score:                  2861.00       Matches:       553
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            44.58%        Indels:        0
DB:                     4             Gaps:          0
```

US-09-759-143-110 (1-3410) x US-09-439-313-113 (1-553)

```
Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 3
US-09-605-785-113
; Sequence 113, Application US/09605785
; Patent No. 6321716
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C16
;  CURRENT APPLICATION NUMBER: US/09/605,785
;  CURRENT FILING DATE:  2000-06-27
;  NUMBER OF SEQ ID NOS: 835
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 113
;   LENGTH: 553
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-605-785-113

Alignment Scores:
Pred. No.:              1.61e-211     Length:        553
Score:                  2861.00       Matches:       553
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            44.58%        Indels:        0
DB:                     4             Gaps:          0

US-09-759-143-110 (1-3410) x US-09-605-785-113 (1-553)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500
```

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/030,607
;       FILING DATE:  25-FEB-1998
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Maki, David J.
;       REGISTRATION NUMBER:  31,392
;       REFERENCE/DOCKET NUMBER:  210121.427C3
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206) 622-4900
;       TELEFAX:  (206) 682-6031
;   INFORMATION FOR SEQ ID NO:  113:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  553 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     ORIGINAL SOURCE:
;       ORGANISM:  Homo sapiens
US-09-030-607-113

Alignment Scores:
Pred. No.:              1.61e-211     Length:          553
Score:                  2861.00       Matches:         553
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            44.58%        Indels:          0
DB:                     4             Gaps:            0

US-09-759-143-110 (1-3410) x US-09-030-607-113 (1-553)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||
```

```
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
US-09-020-956-113

Alignment Scores:
Pred. No.:              1.61e-211     Length:        553
Score:                  2861.00       Matches:       553
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            44.58%        Indels:        0
DB:                     4             Gaps:          0

US-09-759-143-110 (1-3410) x US-09-020-956-113 (1-553)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||
Db        541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 2
US-09-030-607-113
; Sequence 113, Application US/09030607
; Patent No. 6262245
;  GENERAL INFORMATION:
;    APPLICANT:  Xu, Jiangchun
;    APPLICANT:  Dillon, Davin C.
;    TITLE OF INVENTION:  COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS
;    NUMBER OF SEQUENCES:  224
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98104
;    COMPUTER READABLE FORM:
```

GenCore version 5.1.6


```
OM nucleic - protein search, using frame_plus_n2p model

Run on:          June  9, 2003, 22:23:09 ; Search time 29.5 Seconds
                                           (without alignments)
                                           6802.181 Million cell updates/sec

Title:           US-09-759-143-110
Perfect score:   6418
Sequence:        1 gggaaccagcctgcacgcgc..........aaaaaaataaaaaaaaaaaa 3410

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters:      525148

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09759143/runat_05062003_111320_23081/app_query.fasta_1.3591
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=45
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09759143_@CGN_1_1_26_@runat_05062003_111320_23081 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       Issued_Patents_AA:*
                 1:  /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
                 2:  /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
                 3:  /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
                 4:  /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
                 5:  /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                 6:  /cgn2_6/ptodata/2/iaa/backfiles1.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2861	44.6	553	4	US-09-020-956-113	Sequence 113, App
	2	2861	44.6	553	4	US-09-030-607-113	Sequence 113, App
	3	2861	44.6	553	4	US-09-605-785-113	Sequence 113, App
	4	2861	44.6	553	4	US-09-439-313-113	Sequence 113, App
	5	2861	44.6	553	4	US-09-352-616A-113	Sequence 113, App
	6	2861	44.6	553	4	US-09-602-877A-101	Sequence 101, App
	7	2861	44.6	553	4	US-09-232-149A-113	Sequence 113, App
	8	1420.5	22.1	371	4	US-09-605-785-708	Sequence 708, App
	9	1287	20.1	255	4	US-09-071-710-36	Sequence 36, Appl
	10	1287	20.1	255	4	US-09-525-397-36	Sequence 36, Appl
	11	452	7.0	84	4	US-09-439-313-571	Sequence 571, App
	12	426	6.6	123	4	US-09-605-785-706	Sequence 706, App
c	13	395	6.3	762	1	US-08-642-255-120	Sequence 120, App
c	14	395	6.3	762	1	US-08-397-633A-31	Sequence 31, Appl
c	15	393	6.3	761	2	US-08-707-237A-84	Sequence 84, Appl
c	16	391	6.3	762	1	US-08-642-255-114	Sequence 114, App
c	17	391	6.3	762	1	US-08-397-633A-26	Sequence 26, Appl
c	18	374.5	6.0	960	4	US-09-219-849-5	Sequence 5, Appli
c	19	371	6.0	960	4	US-09-219-849-6	Sequence 6, Appli
c	20	370.5	6.0	682	1	US-08-642-255-126	Sequence 126, App
c	21	370.5	6.0	682	1	US-08-397-633A-36	Sequence 36, Appl
	22	365	5.7	761	2	US-08-707-237A-84	Sequence 84, Appl
	23	365	5.7	762	1	US-08-642-255-114	Sequence 114, App
	24	365	5.7	762	1	US-08-397-633A-26	Sequence 26, Appl
c	25	364.5	5.9	1078	3	US-08-963-825-21	Sequence 21, Appl
c	26	364.5	5.9	1078	4	US-09-500-811-21	Sequence 21, Appl
c	27	364.5	5.9	1078	4	US-09-570-573-21	Sequence 21, Appl
c	28	364.5	5.9	1078	4	US-09-548-608-21	Sequence 21, Appl
	29	364	5.7	762	1	US-08-642-255-120	Sequence 120, App
	30	364	5.7	762	1	US-08-397-633A-31	Sequence 31, Appl
c	31	360	5.8	1064	1	US-08-642-255-62	Sequence 62, Appl
c	32	357	5.7	829	1	US-08-642-255-132	Sequence 132, App
c	33	357	5.7	829	1	US-08-397-633A-53	Sequence 53, Appl
c	34	356.5	5.7	837	1	US-08-175-155-68	Sequence 68, Appl
c	35	356.5	5.7	837	1	US-08-477-509B-103	Sequence 103, App
c	36	356.5	5.7	837	1	US-08-642-255-101	Sequence 101, App
c	37	356.5	5.7	837	2	US-08-707-237A-75	Sequence 75, Appl
c	38	356.5	5.7	837	3	US-08-482-085B-103	Sequence 103, App
c	39	356.5	5.7	837	4	US-09-444-791A-103	Sequence 103, App
c	40	356.5	5.7	897	1	US-08-397-633A-50	Sequence 50, Appl
c	41	353.5	5.7	1065	1	US-08-642-255-72	Sequence 72, Appl
c	42	351.5	5.6	1057	3	US-08-931-820-4	Sequence 4, Appli
c	43	351	5.6	822	4	US-09-219-849-49	Sequence 49, Appl
c	44	345	5.5	1057	3	US-08-931-820-1	Sequence 1, Appli
c	45	344.5	5.5	633	1	US-08-642-255-73	Sequence 73, Appl

ALIGNMENTS

```
RESULT 1
US-09-020-956-113
; Sequence 113, Application US/09020956
; Patent No. 6261562
;  GENERAL INFORMATION:
;    APPLICANT:  Xu, Jiangchun
;    APPLICANT:  Dillin, Davin C.
;    TITLE OF INVENTION:  COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS
;    NUMBER OF SEQUENCES:  178
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/020,956
;      FILING DATE:  09-FEB-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.427C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  113:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  553 amino acids
;      TYPE:  amino acid
```

```
Qy       2111 TGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCC 2170
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 TGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCC 191

Qy       2171 TTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGA 2230
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 TTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGA 131

Qy       2231 ATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTG 2290
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTG 71

Qy       2291 AGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTT 2343
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 AGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTT 18
```

```
RESULT 45
BB701488
LOCUS       BB701488                509 bp    mRNA    linear   EST 11-OCT-2001
DEFINITION  BB701488 RIKEN full-length enriched, in vitro fertilized eggs Mus
            musculus cDNA clone 7420429L15 3', mRNA sequence.
ACCESSION   BB701488
VERSION     BB701488.1  GI:16050312
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 509)
  AUTHORS   Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
            Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
            ,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
            Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.
            , Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
            Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
            ,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.
            , Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
            2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .509
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="7420429L15"
                     /clone_lib="RIKEN full-length enriched, in vitro
                     fertilized eggs"
                     /sex="female"
                     /tissue_type="in vitro fertilized eggs"
                     /dev_stage="egg"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. Second strand cDNA was prepared with the
                     primer adapter of sequence [5'
                     GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA
                     was cloned into the XhoI and BamHI sites. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI"
BASE COUNT       68 a    168 c    145 g    128 t
ORIGIN
```

```
  Query Match            10.3%;  Score 352.8;  DB 10;  Length 509;
  Best Local Similarity   83.3%;  Pred. No. 2.1e-37;
  Matches  425;  Conservative    0;  Mismatches   82;  Indels    3;  Gaps    2;

Qy        582 TCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCC 641
              ||| ||||||||| ||||| |||||||| ||| | ||||||| ||| ||  |||||||||
Db          1 TCTGTCTCATCCCGAGGGCTGGCTGGCTGGCA-GACTGCTGTACCCAGACACCAGGCCCC 59

Qy        642 TGGAGCTGGCACTGCTCATCCT--GGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTG 699
              ||||  |||| ||||| || ||  ||    |||   ||||||||| ||||||||||||||
Db         60 TGGATTTGGCCCTGCTGATGCTTGGGAGTGGGGCTGGCTGGACTTTTGTGGCCAGGTGTG 119

Qy        700 CTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCA 759
              ||| |||||| |||||||| | ||||| ||||||||||||||||| |||||||| |||||
Db        120 CTTTACTCCATTGGAGGCCTTACTCTCCGACCTCTTCCGGGACCCAGACCACTGCCGCCA 179

Qy        760 GGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCC 819
               |||| |||| |||| |||||||||||||| | |||||||||||||||||||||| | ||
Db        180 AGCCTTCTCTTTCTACGCCTTCATGATCAGCCGTGGGGGCTGCCTGGGCTACCTCTTACC 239

Qy        820 TGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCT 879
              ||||||||||||||||||||| |  ||||||||||||||||| || ||||| || |||||
Db        240 TGCCATTGACTGGGACACCAGCGTTCTGGCCCCCTACCTGGGTACTCAGGAAGAATGCCT 299

Qy        880 CTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGA 939
              ||||||||| ||||||||||| ||||||| |||| | |||||||| ||| | ||| | ||
Db        300 CTTTGGCCTCCTCACCCTCATTTTCCTCATCTGCATGGCAGCCACTCTGTTTGTGACGGA 359

Qy        940 GGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCC 999
              ||||||||  ||||||||  ||||||| ||||||||| ||| || | | |  || |||||
Db        360 GGAGGCAGTACTGGGCCCACCCGAGCCGGCAGAAGGGTTGTTGGTCTCTGCCGTGTCGCG 419

Qy       1000 CCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCG 1059
              ||  ||||| |||||||  |   || ||||||||||||| |||||  ||||| |||||||
Db        420 CCGATGCTGCCCATGCCACGTTGGCCTGGCTTTCCGGAATCTGGGTACCCTGTTTCCCCG 479

Qy       1060 GCTGCACCAGCTGTGCTGCCGCATGCCCCG 1089
              |||||| |||||||||||||||||||| ||
Db        480 GCTGCAGCAGCTGTGCTGCCGCATGCCTCG 509

Search completed: June  9, 2003, 19:43:37
Job time : 4214 secs
```

```
  Best Local Similarity   78.9%;  Pred. No. 2.3e-38;
  Matches  625;  Conservative    0;  Mismatches  128;  Indels   39;  Gaps   15;

Qy          2 GGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAG 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 GGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAG 75

Qy         62 TGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCA-GCAGGTGTTGAGCATGGGCTGAG 120
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db         76 TGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCATGCAGGTGTTGAGCATGGGCTGAG 135

Qy        121 AAGCTGGACC-GGCACCAAAGGGCTGGCAGAAATGGGCGCCTG-GCTGATTCCTAGGCA- 177
              |||||||||| |||||||||||||||||||||||||||||||| |||||||||||||||
Db        136 AAGCTGGACCTGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGTGCTGATTCCTAGGCAA 195

Qy        178 ----GTTGGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGC 233
                  | ||| |||   |||| |||     |      |    ||||||||||||||||||||
Db        196 GTTGGCGGGCAGCAAGGAGGATGAGGCCGCAGCCTTCTTGGAGCAGAGCCGAGACGAAGC 255

Qy        234 AGTTCTGGAGTGCCT---GAACGGCCCCCTGAGCCCTACCCGCC-TGGCCCACTATGGTC 289
              |||||||||||||||      | |||||||||||||||||||||| |||||||||||||||
Db        256 AGTTCTGGAGTGCCTTGAAACGGGCCCCCTGAGCCCTACCCGCCTTGGCCCACTATGGTC 315

Qy        290 CAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCC--AGCTCTTGCTGGTCA 347
              |||||||||||||||||||||||||||||||||||||||||||    ||||||||||||
Db        316 CAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCCAGCCTCTTGCTGGTCA 375

Qy        348 ACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTC 407
              ||||||||||||||||||||||||||||||| |||           |||||||||||||||
Db        376 ACCTGCTAACCTTTGGCCTGGAGGTGTGTTGGCCGCAGGCATCAACCTATGTGCCGCCTC 435

Qy        408 TGC-----TGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCA-TTG 461
              |||     ||        |||||||||||||||||||||||||||||||||||||| |||
Db        436 TGCTTGCTTGGACACGTGGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTTG 495

Qy        462 GTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACC-ACTGGCGT 520
              ||||||||||||||||||||||||||||||||||||||||||   ||||||| ||||| ||
Db        496 GTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAAGCAGTGACCAACTGGGGT 555

Qy        521 GGACGCTATGG------CCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 574
              |||||||||||      |   ||||   ||||  |||  |      ||||||||||||
Db        556 GGACGCTATGGGGCCGCGCCGCGCCTTTCATTCTTGGGGCACTGTCCTTTGGGCATCCTG 615

Qy        575 CTGAGCCTCTTT--CTCATCCCAAGGGCC------GGCTGGCTAGCAGGGCTGCTGTGCC 626
               |||||||||||  ||||||||  || |         ||||     ||| |    ||||||
Db        616 GTGAGCCTCTTTTCTTCATCCCAGTGGTCCCGGTTGGCTAAGCAGCGGCTTGGTTGTGCC 675

Qy        627 CGGATCCCAGGCCCCTGGA-GCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTC 685
               |   ||  |||||  |||  ||| | |||  |||  |    ||  | |||      || | |
Db        676 TGATCCCAGGGCCCTGGGAGGCTAGGACTTGTCAATCCTGGGTTTGGGGCTGTGCATTCT 735

Qy        686 TGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCG 745
              |||||  |  |  |  | ||   |||   |||||||   ||     ||  |||  |||||||  |||
Db        736 TGTGGGCCAGGTTTCGTCCTTCCCTTGGAGGCCTGTCT---TGGCCTTTTCCGGGGACCG 792

Qy        746 GACCACTGTCGC 757
               | ||     ||||
Db        793 GGCCTTGTTCGC 804
```

```
RESULT 44
AW135465/c
LOCUS       AW135465                 370 bp    mRNA    linear   EST 29-OCT-1999
DEFINITION  UI-H-BI1-aca-e-07-0-UI.s1 NCI_CGAP_Sub3 Homo sapiens cDNA clone
            IMAGE:2713812 3', mRNA sequence.
ACCESSION   AW135465
VERSION     AW135465.1  GI:6139598
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 370)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            The sequence contained an oligo-dT track that was present in the
            oligonucleotide that was used to prime the synthesis of first
            strand cDNA and therefore this may represent a bonafide poly A
            tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:
            NCI-CGAP clone distribution information can be found through the
            I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: M13 Forward
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1..370
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2713812"
                     /clone_lib="NCI_CGAP_Sub3"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The
                     NCI_CGAP_Sub3 library is a subtracted library derived from
                     the  NCI_CGAP_Sub1 library, which is a subtracted library
                     derived from BI.  BI constitutes a mixture of 21
                     normalized or subtracted NCI_CGAP  libraries: NCI_CGAP_Co4
                     , NCI_CGAP_Pr22, NCI_CGAP_Pr28, NCI_CGAP_Co10,
                     NCI_CGAP_Co16, NCI_CGAP_Kid5, NCI_CGAP_Kid12,
                     NCI_CGAP_Kid3,  NCI_CGAP_Kid11, NCI_CGAP_Lym2,
                     NCI_CGAP_Br2, NCI_CGAP_Co8,  NCI_CGAP_CLL1, NCI_CGAP_Lei2,
                     NCI_CGAP_Brn23, NCI_CGAP_Lu5,  NCI_CGAP_Lu24,
                     NCI_CGAP_Lu19, NCI_CGAP_GC4, NCI_CGAP_GC6,
                     NCI_CGAP_Brn25.  These 21 libraries were pooled and a
                     single-stranded DNA  preparation of the resulting mixture
                     was used as a tracer in a subtractive  hybridization with
                     a driver whose composition is detailed below:
                     NCI_CGAP_Kid3 pool 1    LLAM 3334-3337, 3682-3683,
                     3798-3803 (IMAGE CloneIDs 1322376-1323911, 1456008-1456775
                     , 1500552-1502855); NCI_CGAP_Kid5 pool 1    LLAM 3338-3342
                     , 3722-3725, 3776-3778 (IMAGE CloneIDs 1323912-1325831,
                     1471368-1472903, 1492104-1493255); NCI_CGAP_Lu5 pool 1
                     LLAM 3575-3582, 3851-3854 (IMAGE CloneIDs 1414920-1417991,
                     1520904-1522439); NCI_CGAP_GC4 pool 1     LLAM 3164-3167,
                     3716-3720, 3733-3735 (IMAGE CloneIDs 1257096-1258631,
                     1469064-1470983, 1475592-1476743); NCI_CGAP_Pr22 pool 1
                     LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE CloneIDs
                     985608-986759, 1101192-1101959, 1217928-1220615);
                     NCI_CGAP_Co10 pool 1     LLAM 2644-2653, 2871-2872 (IMAGE
                     CloneIDs 1057416-1061255, 1144584-1145351). Subtraction
                     was performed as previously described [Bonaldo, Lennon &
                     Soares (1996): Normalization and Subtraction: Two
                     Approaches To  Facilitate Gene Discovery. Genome Research
                     6, 791-806.
                     TAG_LIB=NCI_CGAP_Pr22
                     TAG_TISSUE=prostate
                     TAG_SEQ=AAGTG"
BASE COUNT       83 a    109 c    101 g     77 t
ORIGIN

  Query Match            10.4%;  Score 353;  DB 10;  Length 370;
  Best Local Similarity  100.0%;  Pred. No. 2.3e-37;
  Matches  353;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1991 CAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCT 2050
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 CAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCT 311

Qy       2051 GCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGC 2110
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGC 251
```

```
                ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db          417 GTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATCGGTCCAGTGCTGGGCCTGGTCTCC 476

Qy          485 GTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCC 544
                ||||| |||||||||||||||||||||||||||||||||| |||||||| ||||| |||||
Db          477 GTCCCACTCCTAGGCTCAGCCAGTGACCACTGGCGTGGGCGCTATGGTCGCCGGAGGCCC 536

Qy          545 T 545
                |
Db          537 T 537


RESULT 42
AA579486/c
LOCUS       AA579486                 388 bp    mRNA    linear   EST 03-SEP-1997
DEFINITION  nf33g07.s1 NCI_CGAP_Pr1 Homo sapiens cDNA clone IMAGE:915612
            similar to contains element MSR1 repetitive element ;, mRNA
            sequence.
ACCESSION   AA579486
VERSION     AA579486.1  GI:2357670
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 388)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index.
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui, M.D.
            , Michael Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 343.
FEATURES             Location/Qualifiers
     source          1..388
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:915612"
                     /clone_lib="NCI_CGAP_Pr1"
                     /sex="Male"
                     /dev_stage="45 years old"
                     /lab_host="DH10B"
                     /note="Vector: pAMP10; Site_1: Not1; Site_2: EcoRI; 1st
                     strand cDNA was primed with oligo(dT)17 on 50 ng of
                     DNAse-treated, total cellular RNA obtained from 5,000-10
                     ,000 microdissected, histologically normal prostate
                     epithelial cells. Double-stranded cDNA was ligated to
                     EcoRI adaptors, 5 cycles of PCR applied to the cDNA with
                     an adaptor-specific primer, and the resulting PCR product
                     subcloned into pAMP10 by the UDG-cloning method (Life
                     Technologies).  Average insert size is 600 bp. NOTE: Not
                     directionally cloned.  This library was constructed by
                     David Krizman."
BASE COUNT       99 a     77 c     99 g    113 t
ORIGIN

  Query Match            10.7%;  Score 366;  DB 9;  Length 388;
  Best Local Similarity  98.7%;  Pred. No. 4.4e-39;
  Matches  369;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         2975 AGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCA 3034
                | |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db          381 AAGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTGCCCCCAACTTTCCCCACCA 322

Qy         3035 GCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAG 3094
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          321 GCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAG 262

Qy         3095 CCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTC 3154
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          261 CCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTC 202

Qy         3155 AGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCG 3214
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          201 AGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCG 142

Qy         3215 TTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCA 3274
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          141 TTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCA 82

Qy         3275 ATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAAAAAA 3334
                ||||||||||||||||||||||||||||||||||||||||||||    |||||||||||||
Db           81 ATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTAACCGTTTAAAAAAAAA 22

Qy         3335 AAAAAAAAAAAAAA 3348
                ||||||||||||||
Db           21 AAAAAAAAAAAAAA 8


RESULT 43
BF680993
LOCUS       BF680993                 906 bp    mRNA    linear   EST 21-DEC-2000
DEFINITION  602156279F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4297170 5',
            mRNA sequence.
ACCESSION   BF680993
VERSION     BF680993.1  GI:11954888
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 906)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONETECH Laboratories, Inc.
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1150  row: n  column: 19
            High quality sequence stop: 564.
FEATURES             Location/Qualifiers
     source          1..906
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4297170"
                     /clone_lib="NIH_MGC_83"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: prostate; Vector: pDNR-LIB (Clontech);
                     Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc
                     ); 5' and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).   Average
                     insert size 1.4 kb (range 0.5-4.0 kb).  14/15 colonies
                     contained inserts  by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA)."
BASE COUNT      134 a    251 c    317 g    203 t      1 others
ORIGIN

  Query Match            10.5%;  Score 359.2;  DB 12;  Length 906;
```

```
              prepared and sequenced in Mouse Genome Encyclopedia
              Project of Genome Exploration Research Group in Riken
              Genomic Sciences Center and Genome Science Laboratory in
              RIKEN. Division of Experimental Animal Research in Riken
              contributed to prepare mouse tissues. 1st strand cDNA was
              primed with a primer [5'
              GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
              prepared by using trehalose thermo-activated reverse
              transcriptase and subsequently enriched for full-length by
              cap-trapper. cDNA went through one round of normalization
              to Rot = 20.0 and subtraction to Rot = 370.0. Second
              strand cDNA was prepared with the primer adapter of
              sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC
              3'].  cDNA was cleaved with XhoI and BamHI. Vector: a
              modified pBluescript KS(+) after bulk excision from
              Lambda FLC I."
BASE COUNT      120 a    182 c    214 g    134 t
ORIGIN

  Query Match             11.2%;  Score 381.2;  DB 10;  Length 650;
  Best Local Similarity   79.2%;  Pred. No. 3.4e-41;
  Matches  494;  Conservative    0;  Mismatches  118;  Indels   12;  Gaps    3;

Qy          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
Db         37 GGGTCCGAGCTCGCACGCGCCAGCCCCAGGTGACAGCCGCACGCC-GGGCCAGGATCTGA 95

Qy         61 GTGATGAGACGTGTCCCCACTGAGGT--GCCCCACAGCAGCAGGTGTTGAGCATGGGCTG 118
Db         96 CCGACGAGATGTGTCCCCATCAAGCAAGGCACTAGATGGTGACGTGTTTAGCGTGGGACG 155

Qy        119 AGAAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAG 178
Db        156 AGATGCTGAATTGGCACTAAAGGGCTGGCAGAAATGGGAACCTGGCTGCACCCTAGGAGG 215

Qy        179 TTGGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTC 238
Db        216 TTAGTGCTAGTGAGGAGGAGAAGCCAC---------GGCAGGGCTGACTCAAAGCAGCTG 266

Qy        239 TGGAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTG 298
Db        267 TGGAGTATGTGAGTAGCCCCTGGAACCCTACCTGCCCTGTCCATCATGATCCAGAGGCTG 326

Qy        299 TGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACC 358
Db        327 TGGGCCAGCCGTCTGCTACGGCACCGGAAAGCTCAGCTCCTGCTGGTCAACCTGCTCACC 386

Qy        359 TTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAA 418
Db        387 TTTGGCCTGGAGGTGTGCCTGGCTGCCGGCATTACCTATGTGCCACCCCTTCTGCTGGAA 446

Qy        419 GTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTG 478
Db        447 GTCGGGGTGGAGGAGAAATTCATGACCATGGTGTTGGGCATTGGCCCAGTGCTAGGCCTG 506

Qy        479 GTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGC 538
Db        507 GTTTCTGTTCCACTCCTAGGCTCAGCCAGTGACCAGTGGCGTGGGCGCTATGGCCGCCGG 566

Qy        539 CGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGG 598
Db        567 AGACCCTTTATCTGGGCTTTGTCCCTGGGTGTCCTGCTAAGCCTCTTTCTCATCCCGAGG 626

Qy        599 GCCGGCTGGCTAGCAGGGCTGCTG 622
Db        627 GCTGGCTGGCTGGCAGGACTGCTG 650

RESULT 41
AW787124
LOCUS       AW787124                 537 bp    mRNA    linear   EST 09-JUL-2000
DEFINITION  120845 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   AW787124
VERSION     AW787124.1  GI:7843900
KEYWORDS    EST.
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 537)
  AUTHORS   Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
            Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
            and Keele,J.W.
  TITLE     Design and use of two pooled tissue normalized cDNA libraries for
            EST discovery in swine
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 45  row: C  column: 12
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
                     /clone_lib="MARC 1PIG"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
                     Library made from pooled tissue from day 11, 13, 15, 20,
                     and 30 embryos."
BASE COUNT       96 a    161 c    186 g     94 t
ORIGIN

  Query Match             10.8%;  Score 369.4;  DB 10;  Length 537;
  Best Local Similarity   85.0%;  Pred. No. 1.3e-39;
  Matches  460;  Conservative    0;  Mismatches   76;  Indels    5;  Gaps    4;

Qy          5 ACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTGA 64
Db          2 ACGAGCTTGCACGCGCCGGCCTCGGGTGACAGCCTCCCTCCTCCGCCAGGATCTGAGCCA 61

Qy         65 TGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAGC 124
Db         62 TGAGTTATGT-CCCACCGAAGCACCCTACAGTGACACGCGGTTAGCAT-GGCTGAGAAGC 119

Qy        125 TGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGCG 184
Db        120 TAGACCTGCA-TGGAGGGCTGGCAGAAATGGGTGCCTGGCCGAGTCCTAGGCAGTTGGTG 178

Qy        185 GCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAGT 244
Db        179 GCAGCAAGGAGGAGAGGCCGGGGCGTCGGGGGCAAGGCCGAGCTGGAGCAGTTCCGAAGT 238

Qy        245 GCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGTG 304
Db        239 GCATGAACGG--GCCAGAGCCCTACCCACCCGGCCCATCATGGTCCAGAGACTGTGGATG 296

Qy        305 AGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGGC 364
Db        297 AGCCGCCTGCTGCGGCATCGGAAAGCCCAGCTCCTGCTGGTTAACCTGCTGACGTTCGGC 356

Qy        365 CTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGG 424
Db        357 CTGGAGGTGTGCCTGGCCGCAGGCATCACCTACGTGCCACCCCTGCTGTTGGAAGTGGGG 416

Qy        425 GTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGT 484
```

```
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11137  row: n  column: 24
            High quality sequence start: 2
            High quality sequence stop: 630.
FEATURES             Location/Qualifiers
     source          1. .630
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5050583"
                     /clone_lib="NCI_CGAP_Li9"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.9 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      114 a    180 c    199 g    137 t
ORIGIN

  Query Match              11.4%;  Score 389.4;  DB 13;  Length 630;
  Best Local Similarity    79.4%;  Pred. No. 2.9e-42;
  Matches  490;  Conservative    0;  Mismatches  116;  Indels   11;  Gaps    2;

Qy         47 CGGCCAGGATCTGAGTGATGAGACGTGTCCCCACTGAG--GTGCCCCACAGCAGCAGGTG 104
              ||| ||||||||||  || |||| |||||||||   ||    || | | |     | |||
Db         21 CGGTCAGGATCTGACCGACGAGATGTGTCCCCATCAAGCAAGGCACTAGATGGTGACGTG 80

Qy        105 TTGAGCATGGGCTGAGAAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGC 164
              || ||| ||||  |||| |||| |  ||||| ||||||||||||||||||||  ||||||
Db         81 TTTAGCGTGGGACGAGATGCTGAATTGGCACTAAAGGGCTGGCAGAAATGGGAACCTGGC 140

Qy        165 TGATTCCTAGGCAGTTGGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCG 224
              ||   ||||||  ||| | |  ||  ||||||||| ||| |          |||| || |
Db        141 TGCACCCTAGGAGGTTAGTGCTAGTGAGGAGGAGAAGCCAC---------GGCAGGGCTG 191

Qy        225 AGACGAAGCAGTTCTGGAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTA 284
              |  | |||||| | ||||||   |||   |||||  || ||||||| |||  | |||  |
Db        192 ACTCAAAGCAGCTGTGGAGTATGTGAGTAGCCCCTGGAACCCTACCTGCCCTGTCCATCA 251

Qy        285 TGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGG 344
              || |||||||||||||||  ||||| ||||| ||||||||||||| ||||||| ||||||
Db        252 TGATCCAGAGGCTGTGGGCCAGCCGTCTGCTACGGCACCGGAAAGCTCAGCTCCTGCTGG 311

Qy        345 TCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGC 404
              |||||||||| ||||||||||||||||||||  |||| || ||||| ||||||||||| |
Db        312 TCAACCTGCTCACCTTTGGCCTGGAGGTGTGCCTGGCTGCCGGCATTACCTATGTGCCAC 371

Qy        405 CTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTC 464
              | || ||||||||||| ||||| |||||||| ||||||||||||||| |||||||||| |
Db        372 CCCTTCTGCTGGAAGTCGGGGTGGAGGAGAAATTCATGACCATGGTGTTGGGCATTGGCC 431

Qy        465 CAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGAC 524
              ||||||| |||||||| | ||| || ||||||||||||||||||||||| |||||||| |
Db        432 CAGTGCTAGGCCTGGTTTCTGTTCCACTCCTAGGCTCAGCCAGTGACCAGTGGCGTGGGC 491

Qy        525 GCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCT 584
              |||||||||||||  | ||||| ||||||||  ||||| ||||  ||||||| |||||||
Db        492 GCTATGGCCGCCGGAGACCCTTTATCTGGGCTTTGTCCCTGGGTGTCCTGCTAAGCCTCT 551

Qy        585 TTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGG 644
              |||||||||| ||||| |||||||| ||||| ||||||| ||| ||  ||||||||||||
Db        552 TTCTCATCCCGAGGGCTGGCTGGCTGGCAGGACTGCTGTACCCAGACACCAGGCCCCTGG 611

Qy        645 AGCTGGCACTGCTCATC 661
              ||  ||| ||||| |||
Db        612 AGTAGGCCCTGCTGATC 628


RESULT 40
BB627844
LOCUS       BB627844                 650 bp    mRNA    linear   EST 26-OCT-2001
DEFINITION  BB627844 RIKEN full-length enriched, adult male urinary bladder Mus
            musculus cDNA clone 9530042D02 5', mRNA sequence.
ACCESSION   BB627844
VERSION     BB627844.1  GI:16465416
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 650)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
            , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
            ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="9530042D02"
                     /clone_lib="RIKEN full-length enriched, adult male urinary
                     bladder"
                     /sex="male"
                     /tissue_type="urinary bladder"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
```

```
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV2&t2=QV2-NT0147-
            191100-482-h07&t3=2000-11-19&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 418.
FEATURES             Location/Qualifiers
     source          1..418
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="NT0147"
                     /dev_stage="Adult"
                     /note="Organ: nervous_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT      104 a    122 c    127 g     65 t
ORIGIN

  Query Match             11.6%;  Score 395.8;  DB 12;  Length 418;
  Best Local Similarity   99.5%;  Pred. No. 5.1e-43;
  Matches  397;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1730 GTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCC 1789
              ||| |||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 GTTTCGGGACGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCC 352

Qy       1790 CAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCC 1849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 CAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCC 292

Qy       1850 TATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTA 1909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 TATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTA 232

Qy       1910 TTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGG 1969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 TTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGG 172

Qy       1970 AGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGC 2029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 AGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGC 112

Qy       2030 TGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGC 2089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        111 TGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGC 52

Qy       2090 TGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCC 2128
              |||||||||||||||||||||||||||||||||||||||
Db         51 TGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCC 13

RESULT 38
AF109303
LOCUS       AF109303                 430 bp    mRNA    linear   EST 28-JAN-2000
DEFINITION  AF109303 Homo sapiens prostate adult Homo sapiens cDNA clone IPCA-8
            , mRNA sequence.
ACCESSION   AF109303
VERSION     AF109303.1  GI:6782698
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 430)
  AUTHORS   Walker,M.G., Volkmuth,W., Sprinzak,E., Hodgson,D. and Klingler,T.
  TITLE     Prediction of gene function by genome-scale expression analysis:
            prostate cancer-associated genes
  JOURNAL   Genome Res. 9 (12), 1198-1203 (1999)
  MEDLINE   20082966
COMMENT     Contact: Walker MG
            Incyte Pharmaceuticals
            3174 Porter Drive, Palo Alto, CA 94304, USA
            co-expressed with known prostate-cancer genes.
FEATURES             Location/Qualifiers
     source          1..430
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IPCA-8"
                     /clone_lib="Homo sapiens prostate adult"
                     /tissue_type="prostate"
                     /dev_stage="adult"
                     /note="multiple clone assembly from multiple libraries and
                     donors"
BASE COUNT       81 a    119 c    150 g     80 t
ORIGIN

  Query Match             11.5%;  Score 391;  DB 9;  Length 430;
  Best Local Similarity   97.0%;  Pred. No. 2.1e-42;
  Matches  420;  Conservative    0;  Mismatches   10;  Indels    3;  Gaps    2;

Qy         33 ACAGCCGCGCGCCTCGGCCAGGATCTGAGTGATGAGACGTGTCCCCACTGAGGTGCCCCA 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACAGCCGCGCGCCTCGGCCAGGATCTGAGTGATGAGACGTGTCCCCACTGAGGTGCCCCA 60

Qy         93 CAGCAGCAGGTGTTGAGCATGGGCTGAGAAGCTGGACCGGCACCAAAGGGCTGGCAGAAA 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGCAGCAGGTGTTGAGCATGGGCTGAGAAGCTGGACCGGCACCAAAGGGCTGGCAGAAA 120

Qy        153 TGGGCGCCTGGCTGATTCCTAGGCAGTTGGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCT 212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGGGCGCCTGGCTGATTCCTAGGCAGTTGGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCT 180

Qy        213 GGAGCAGAGCCGAGACGAAGCAGTTCTGGAGTGCCTGAACGGCCCCCTGAGCCCTACCCG 272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAGCAGAGCCGAGACGAAGCAGTTCTGGAGTGCCTGAACGGCCCCCTGAGCCCTACCCG 240

Qy        273 CCTGGCCCACTATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCC 332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGGCCCACTATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCC 300

Qy        333 AGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCA 392
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| |  |||
Db        301 AGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCA-GATTCA 359

Qy        393 CCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGC 452
              |||||||||||||||||      | | |||||||||||||||||||||||||||||||||
Db        360 CCTATGTGCCGCCTCTG--TGCTGGAATGGGGGTAGAGGAGAAGTTCATGACCATGGTGC 417

Qy        453 TGGGCATTGGTCC 465
              ||||  |||||||
Db        418 TGGGATTTGGTCC 430

RESULT 39
BI145201
LOCUS       BI145201                 630 bp    mRNA    linear   EST 05-JUL-2001
DEFINITION  602909395F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:5050583 5',
            mRNA sequence.
ACCESSION   BI145201
VERSION     BI145201.1  GI:14605202
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 630)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
```

```
Db        177 AAAGGGCTGGCAGAAATGGGAACCTGGCTGCACCCTAGGAGGTTAGTGCTAGTGAGGAGG 236

Qy        197 AGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAGTGCCTGAACGGCC 256
               ||| ||| |          |||| || ||  | |||||| | ||||||    |||    |||
Db        237 AGAAGCCAC---------GGCAGGGCTGACTCAAAGCAGCTGTGGAGTATGTGAGTAGCC 287

Qy        257 CCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTG 316
              ||  || ||||||| |||   | |||   || ||||||||||||||  ||||| |||||
Db        288 CCTGGAACCCTACCTGCCCTGTCCATCATGATCCAGAGGCTGTGGGCCAGCCGTCTGCTA 347

Qy        317 CGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGT 376
              ||||||||||||| |||||| ||||||||||||||| ||||||||||||||||||||
Db        348 CGGCACCGGAAAGCTCAGCTCCTGCTGGTCAACCTGCTCACCTTTGGCCTGGAGGTGTGC 407

Qy        377 TTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGA-AGTGGGGTAGAGGAGAA 435
                |||| || ||||| |||||||||| || || |||||||| ||| ||||| ||||||||
Db        408 CTGGCTGCCGGCATTACCTATGTGCCACCCCTTCTGCTGGACAGTCGGGGTGGAGGAGAA 467

Qy        436 GTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGG-TCTGTGTCCCGCTCC 494
                ||||||||||||||| |||||||||| |||||||| ||||||| | | ||| || ||||
Db        468 ATTCATGACCATGGTGTTGGGCATTGGCCCAGTGCTAGGCCTGGTTTTCTGTTCCACTCC 527

Qy        495 TAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGG 554
              ||||||||||||||||||| ||||||| |||||||||||||     | ||||| |||||||
Db        528 TAGGCTCAGCCAGTGACCAGTGGCGTGGGCGCTATGGCCGCCGGAGACCCTTTATCTGGG 587

Qy        555 CACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAG 614
              |     |||| ||||  ||||||| |||||||||||||||| ||||| | ||||||| ||||
Db        588 CTTCGTCCCTGGGTGTCCTGCTAAGCCTCTTTCTCATCCCGAGGGCTGTCTGGCTGGCAG 647

Qy        615 GGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGC 674
               | ||||||| ||| ||   ||||| |||||||| |||| ||| | ||| |||  |||||||
Db        648 GACTGCTGTACCCAGAGACCAGG-CCCTGGAGTTGGCCCTGGTGATCTTGGAAGTGGGGC 706

Qy        675 TGCTGGACTTCTGTGGCCAGGTGTGC-TTCACTCCACTGGAGGCCCTGCTCTCTGACCTC 733
              || ||||| |  |||||||| |||| || |||||| |||||||||  |    | |    |
Db        707 TGGTGGACCTTCGTGGCCAGTGGTGCTTTAACTCCATTGGAGGCCCTAATTCTTCAGACC 766

Qy        734 TTCCGGGACCCGGACCACTGTCGCCAG 760
              || | ||  ||  | ||||| || ||
Db        767 TTTCCGGGACCCAAACACTGCGGCAAG 793

RESULT 36
AW175665/c
LOCUS       AW175665              412 bp    mRNA    linear   EST 16-NOV-1999
DEFINITION  RC3-BT0046-310899-002-F07 BT0046 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW175665
VERSION     AW175665.1  GI:6441806
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 412)
  AUTHORS   HCGP http://www.ludwig.org.br/ORESTES.
  TITLE     The FAPESP/LICR Human Cancer Genome Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC3&t2=RC3-BT0046-
            310899-002-F07&t3=1999-08-31&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 411.
FEATURES             Location/Qualifiers
     source          1..412
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="BT0046"
                     /dev_stage="Adult"
                     /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No. 196
                     ,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
BASE COUNT      106 a    117 c    128 g     61 t
ORIGIN

  Query Match            11.7%;  Score 398.6;  DB 10;  Length 412;
  Best Local Similarity   99.0%;  Pred. No. 2.2e-43;
  Matches  401;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1795 GGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATAT 1854
              | | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        405 GCCTCCTTCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATAT 346

Qy       1855 GGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGA 1914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        345 GGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGA 286

Qy       1915 CAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGC 1974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        285 CAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGC 226

Qy       1975 CTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCC 2034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        225 CTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCC 166

Qy       2035 GCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGG 2094
              |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        165 GCCAGTTTCTGTTGCTGCCAAAGTAATGCGGCTCTCTGCTGCCACCCTGTGCTGCTGAGG 106

Qy       2095 TGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGC 2154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        105 TGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGC 46

Qy       2155 TGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATA 2199
              |||||||||||||||||||||||||||||||||||||||||||||
Db         45 TGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATA 1

RESULT 37
BF922235/c
LOCUS       BF922235              418 bp    mRNA    linear   EST 19-JAN-2001
DEFINITION  QV2-NT0147-191100-482-h07 NT0147 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF922235
VERSION     BF922235.1  GI:12318123
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 418)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
```

```
          Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
           Please visit our web site (http://genome.gsc.riken.go.jp) for
          further details.
           e mouse tissues.
FEATURES             Location/Qualifiers
     source          1..692
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="9530024M01"
                     /clone_lib="RIKEN full-length enriched, adult male urinary
                     bladder"
                     /sex="male"
                     /tissue_type="urinary bladder"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 20.0 and subtraction to Rot = 370.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC
                     3']. cDNA was cleaved with XhoI and BamHI. Vector: a
                     modified pBluescript  KS(+) after bulk excision from
                     Lambda FLC I."
BASE COUNT      127 a    197 c    224 g    143 t      1 others
ORIGIN

  Query Match             12.0%;  Score 409.4;  DB 10;  Length 692;
  Best Local Similarity   79.1%;  Pred. No. 6.5e-45;
  Matches  527;  Conservative    0;  Mismatches  127;  Indels   12;  Gaps    3;

Qy          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
              |||  | |||  ||||||||  || ||  ||||||||||| ||||  ||||||||||||
Db         37 GGGTCCGAGCTCGCACGCGCCAGCCCCAGGTGACAGCCGCACGCC-GGGCCAGGATCTGA 95

Qy         61 GTGATGAGACGTGTCCCCACTGAGGT--GCCCCACAGCAGCAGGTGTTGAGCATGGGCTG 118
                || ||||  |||||||||    ||       || | | |      | ||||| ||| ||||   |
Db         96 CCGACGAGATGTGTCCCCATCAAGCAAGGCACTAGATGGTGACGTGTTTAGCGTGGGACG 155

Qy        119 AGAAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAG 178
              ||| |||| |  ||||| ||||||||:||||||||  ||||||||   ||||||  |
Db        156 AGATGCTGAATTGGCACTAAAGGGCTGGCAGAAATGGGAACCTGGCTGCACCCTAGGAGG 215

Qy        179 TTGGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTC 238
              || |  |  ||  |||||||| ||| |          |||| || ||  | |||||| |
Db        216 TTAGTGCTAGTGAGGAGGAGAAGCCAC---------GGCAGGGCTGACTCAAAGCAGCTG 266

Qy        239 TGGAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTG 298
              |||||  |||  |||||  || ||||||| |||  | |||  ||| |||||||||||
Db        267 TGGAGTATGTGAGTAGCCCCCTGGAACCCTACCTGCCCTGTCCATCATGATCCAGAGGCTG 326

Qy        299 TGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACC 358
              ||||  ||||| ||||| ||||||||||||| |||||| ||||||||||||||| |||
Db        327 TGGGCCAGCCGTCTGCTACGGCACCGGAAAGCTCAGCTCCTGCTGGTCAACCTGCTCACC 386

Qy        359 TTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAA 418
              ||||||||||||||||  |||| || ||||| |||||||||| || || |||||||||
Db        387 TTTGGCCTGGAGGTGTGCCTGGCTGCCGGCATTACCTATGTGCCACCCCTTCTGCTGGAA 446

Qy        419 GTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTG 478
              || ||||| |||||||| ||||||||||||||| |||||||||| |||||||| |'||||
Db        447 GTCGGGGTGGAGGAGAAATTCATGACCATGGTGTTGGGCATTGGCCCAGTGCTAGGCCTG 506

Qy        479 GTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGC 538
              || | ||| || ||||||||||||||||||||||| |||||||| |||||||||||||
Db        507 GTTTCTGTTCCACTCCTAGGCTCAGCCAGTGACCAGTGGCGTGGGCGCTATGGCCGCCGG 566

Qy        539 CGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGG 598
                | ||||| ||||||||  ||||| ||||  ||||||| |||||||||||||||| |||
Db        567 AGACCCTTTATCTGGGCTTTGTCCCTGGGTGTCCTGCTAAGCCTCTTTCTCATCCCGAGG 626

Qy        599 GCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTC 658
              || |||||||| ||||| ||||||| ||| ||   ||| |||||||||| |||| |||||
Db        627 GCTGGCTGGCTGGCAGGACTGCTGTACCCAGACACCANGCCCCTGGAGTTGGCCCTGCTG 686

Qy        659 ATCCTG 664
              ||| ||
Db        687 ATCTTG 692


RESULT 35
BG246497
LOCUS       BG246497                 844 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  602360526F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4488895 5',
            mRNA sequence.
ACCESSION   BG246497
VERSION     BG246497.1  GI:12756312
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 844)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10336  row: c  column: 08
            High quality sequence stop: 655.
FEATURES             Location/Qualifiers
     source          1..844
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4488895"
                     /clone_lib="NCI_CGAP_Mam1"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="10 months, virgin"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
BASE COUNT      152 a    251 c    262 g    179 t
ORIGIN

  Query Match             11.7%;  Score 400.2;  DB 12;  Length 844;
  Best Local Similarity   77.4%;  Pred. No. 9.5e-44;
  Matches  578;  Conservative    0;  Mismatches  153;  Indels   16;  Gaps    7;

Qy         19 GCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTGATGAGACGTGTCCCC 78
              ||  || || ||||||||||||| ||||  |||||||||||||   || |||| ||||||||
Db         58 GCCAGCCCCAGGTGACAGCCGCACGCC-GGGCCAGGATCTGACCGACGAGATGTGTCCCC 116

Qy         79 ACTGAGGT--GCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAGCTGGACCGGCACC 136
              |    ||      || | | |      | ||||| ||| ||||  ||||| ||||  |  |||||
Db        117 ATCAAGCAAGGCACTAGATGGTGACGTGTTTAGCGTGGGACGAGATGCTGAATTGGCACT 176

Qy        137 AAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGCGGCAGCAAGGAGG 196
              |||||||||||||||||||  ||||||||   ||||||   ||| |  |   ||  ||||||
```

```
VERSION     BF789072.1  GI:12094108
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 759)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9810  row: 1  column: 07
            High quality sequence start: 3
            High quality sequence stop: 756.
FEATURES             Location/Qualifiers
     source          1..759
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4222998"
                     /clone_lib="NCI_CGAP_Kid14"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library. |"
BASE COUNT      131 a    218 c    243 g    167 t
ORIGIN

  Query Match             12.3%;  Score 419.2;  DB 12;  Length 759;
  Best Local Similarity   79.4%;  Pred. No. 3.2e-46;
  Matches  575;  Conservative    0;  Mismatches  133;  Indels   16;  Gaps    6;

Qy          9 GCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTGATGAG 68
                ||  |||||||  || ||  ||||||||||| ||||  ||||||||||||  || |||
Db         16 GCTCGCACGCGCCAGCCCCAGGTGACAGCCGCACGCC-GGGCCAGGATCTGACCGACGAG 74

Qy         69 ACGTGTCCCCACTGAGGT--GCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAGCTG 126
               | |||||||||    ||      || | | |     | ||||| ||| ||||  |||| |||
Db         75 ATGTGTCCCCATCAAGCAAGGCACTAGATGGTGACGTGTTTAGCGTGGGACGAGATGCT- 133

Qy        127 GACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGCGGC 186
              ||  ||||| |||||||||||||||||||  ||||||||    ||||||  ||| | |
Db        134 GAATGGCACTAAAGGGCTGGCAGAAATGGGAACCTGGCTGCACCCTAGGAGGTTAGTGCT 193

Qy        187 AGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAGTGC 246
              ||  |||||||||| ||| |              |||| || ||  | |||||| | ||||||
Db        194 AGTGAGGAGGAGAAGCCACG---------GCAGGGCTGACTCAAAGCAGCTGTGGAGTAT 244

Qy        247 CTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGTGAG 306
               |||  |||||  || || ||||||| ||  | |||  |||  ||||||||||||||||  ||
Db        245 GTGAGTAGCCCCTGGAACCCTACCTGCCCTGTCCATCATGATCCAGAGGCTGTGGGCCAG 304

Qy        307 CCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCT 366
              ||| ||||| ||||||||||||| |||||| ||||||||||||||||| |||||||||||
Db        305 CCGTCTGCTACGGCACCGGAAAGCTCAGCTCCTGCTGGTCAACCTGCTCACCTTTGGCCT 364

Qy        367 GGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGT 426
              |||||||||  |||| || ||||| ||||||||||| || || |||||||||||| |||||
Db        365 GGAGGTGTGCCTGGCTGCCGGCATTACCTATGTGCCACCCCTTCTGCTGGAAGTCGGGGT 424

Qy        427 AGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGT 486
               |||||||| ||||||||||||||| |||||||||| |||||||| |||||||| | |||
Db        425 GGAGGAGAAATTCATGACCATGGTGTTGGGCATTGGCCCAGTGCTAGGCCTGGTTTCTGT 484

Qy        487 CCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTT 546
               || ||||||||||||||||||||| |||||||| |||||||||| ||   | |||||
Db        485 TCCACTCCTAGGCTCAGCCAGTGACCAGTGGCGTGGGCGCTATGGCCG-CGGAGACCCTT 543

Qy        547 CATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTG 606
               ||||||||  ||||| ||||  ||||||| |||||||||||||||||||  ||||| ||||
Db        544 TATCTGGGCTTTGTCCCTGGGTGTCCTGCTAAGCCTCTTTCTCATCCCGAGGGCTGGCTG 603

Qy        607 GCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGG 666
              ||| ||||| ||||||| ||| ||    ||||||||||||| |||| ||||| ||| ||||
Db        604 GCTGGCAGGACTGCTGTACCCAGACACCAGGCCCCTGGAGTTGGCCCTGCTGATCTTGGG 663

Qy        667 CGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTC 726
                |||||||||||||||   |    | |||||||||| |||||| |||||| | | |||||
Db        664 AGTGGGGCTGCTGGACT--TTGTGGCAGGTGTGCTTTACTCCATTGGAGGGCTTACTCTC 721

Qy        727 TGAC 730
               |||
Db        722 CGAC 725


RESULT 34
BB627667
LOCUS       BB627667                 692 bp    mRNA    linear   EST 26-OCT-2001
DEFINITION  BB627667 RIKEN full-length enriched, adult male urinary bladder Mus
            musculus cDNA clone 9530024M01 5', mRNA sequence.
ACCESSION   BB627667
VERSION     BB627667.1  GI:16465271
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 692)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
            , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
            ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
             Computational Analysis of Full-Length Mouse cDNAs Compared with
```

```
Db        667 CCCTCTACCACCGTGAGAAGCCGGTGTCC 695

RESULT 32
BB610495
LOCUS       BB610495                 700 bp    mRNA    linear   EST 26-OCT-2001
DEFINITION  BB610495 RIKEN full-length enriched, adult male stomach Mus
            musculus cDNA clone 2210413P12 5', mRNA sequence.
ACCESSION   BB610495
VERSION     BB610495.1  GI:16452013
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 700)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
            , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
            ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1..700
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="2210413P12"
                     /clone_lib="RIKEN full-length enriched, adult male
                     stomach"
                     /sex="male"
                     /tissue_type="stomach"
                     /dev_stage="adult"
                     /lab_host="SOLR"
                     /note="Site_1: XhoI; Site_2: SstI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAGCGGCCGCAACTCGAGTTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. Second strand cDNA was prepared with the
                     primer adapter of sequence [5'
                     GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA
                     was cleaved with XhoI and SstI."
BASE COUNT      125 a    201 c    206 g    168 t
ORIGIN

  Query Match             12.3%;  Score 419.6;  DB 10;  Length 700;
  Best Local Similarity   78.8%;  Pred. No. 2.9e-46;
  Matches  547;  Conservative    0;  Mismatches  124;  Indels   23;  Gaps    3;

Qy       1458 TGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGC 1517
              ||||||||||||| ||||| || |||||||||||||||| ||||||||||||||||||
Db          2 TGCAGATCCTGCCTTACACGCTCGCCTCCCTCTACCACCGTGAGAAGCAGGTGTTCCTGC 61

Qy       1518 CCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCC 1577
              |||||||||||||||| ||||||||    ||| |||||||||||| ||  |||||||||
Db         62 CCAAATACCGAGGGGACGCTGGAGGTAGCAGCGGTGAGGACAGCCAGACAACCAGCTTCT 121

Qy       1578 TGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTG 1637
              ||||||||||||||| ||||||| ||||||  |||||||||||||||    ||||  |||| |
Db        122 TGCCAGGCCCTAAGCCAGGAGCTCTCTTCCCCAATGGACACGTGGGCTCTGGCAGCAGCG 181

Qy       1638 GCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTG 1697
              || | ||  | |||||||| || ||||| |||||||||||||| ||||||||| | || |
Db        182 GCATCCTGGCCCCTCCACCTGCACTCTGTGGGGCCTCTGCCTGCGATGTCTCCATGCGAG 241

Qy       1698 TGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACC 1757
              |||||||||||||||   | ||||||||||| ||| |||| |||||||| ||||||||||
Db        242 TGGTGGTGGGTGAGCCACCTGAGGCCAGGGTTGTTACGGGACGGGGCATTTGCCTGGACC 301

Qy       1758 TCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCT 1817
              | ||||| ||||| |||||||| |||||||||||||||| || |||||||| |||||||
Db        302 TGGCCATTCTGGACAGTGCCTTTCTGCTGTCCCAGGTGGCTCCGTCCCTGTTCATGGGCT 361

Qy       1818 CCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTC 1877
              |||||||||||| ||||| ||||||||||||||||||| || |||||| ||||||||
Db        362 CCATTGTCCAGCTGAGCCACTCTGTCACTGCCTATATGGTATCAGCTGCAGGCTTGGGTC 421

Qy       1878 TGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACT 1937
              |||||||||||||||||||||||||||||| |||||||||| ||||||||||||||||||
Db        422 TGGTCGCCATTTACTTTGCTACACAGGTAGTGTTTGACAAGAACGACTTGGCCAAATACT 481

Qy       1938 CAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCC 1997
              ||| ||||||   |    ||       |     ||||| ||||||||  || || ||||
Db        482 CAGTGTAGAATTGTGTAAG--GCATCAAAGAGAGGGTCTGCCTCATGGGTTCTCAGC--- 536

Qy       1998 CCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAG 2057
                           ||  | ||||||| | ||||||| |            |||  ||| ||
Db        537 -------------CCTTAGAGGGCTGCAGAGCTGGCCTCTCCAGGTCTTTGTCGCCTAA- 582

Qy       2058 TAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2117
                  |||||||||||||||||||||   |      |||||||  | | ||||||| | || ||
Db        583 ----GTGGCTCTCTGCTGCCACCCTAAGGCAATGAGGTGTATTGTTGCACAGATAGGAGC 638

Qy       2118 TGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAG 2151
                | || | ||    ||||  |   |||   ||| |
Db        639 CAGAGCTTTCCGGGGCTCTGGCTTCAGAGTCTGG 672


RESULT 33
BF789072
LOCUS       BF789072                 759 bp    mRNA    linear   EST 12-JAN-2001
DEFINITION  602104930F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4222998
            5', mRNA sequence.
ACCESSION   BF789072
```

```
                     /clone="IMAGE:3278742"
                     /clone_lib="NCI_CGAP_Lu24"
                     /tissue_type="carcinoid"
                     /lab_host="DH10B"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Plasmid DNA from the normalized
                     library NCI_CGAP_Lu5 was prepared, and ss circles were
                     made in vitro. Following HAP purification, this DNA was
                     used as tracer in a subtractive hybridization reaction.
                     The driver was PCR-amplified cDNAs from a pool of 5,000
                     clones made from the same library (cloneIDs
                     1414920-1417991 and 1520904-1522439). Subtraction by Bento
                     Soares and M. Fatima Bonaldo. "
BASE COUNT       96 a    142 c     96 g    114 t
ORIGIN

  Query Match             12.8%;  Score 435.4;  DB 10;  Length 448;
  Best Local Similarity   99.6%;  Pred. No. 3e-48;
  Matches  447;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       2626 TAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGG 2685
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGG 60

Qy       2686 GATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATC 2745
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         61 GATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCT-ATCATTGCCAGAATC 119

Qy       2746 TTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTC 2805
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 TTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTC 179

Qy       2806 ATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAG 2865
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 ATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAG 239

Qy       2866 GGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGG 2925
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 GGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGG 299

Qy       2926 TTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCC 2985
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 TTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCC 359

Qy       2986 AAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACC 3045
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 AAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACC 419

Qy       3046 CTGTTTGGAGCTACTGCAGGACCAGAAGC 3074
              |||| ||||||||||||||||||||||||
Db        420 CTGTCTGGAGCTACTGCAGGACCAGAAGC 448

RESULT 31
BF581244
LOCUS       BF581244                 700 bp    mRNA    linear   EST 12-DEC-2000
DEFINITION  602100464F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4220415 5',
            mRNA sequence.
ACCESSION   BF581244
VERSION     BF581244.1  GI:11654956
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 700)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9803  row: p  column: 16
            High quality sequence start: 2
            High quality sequence stop: 680.
FEATURES             Location/Qualifiers
     source          1..700
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4220415"
                     /clone_lib="NCI_CGAP_Co24"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.6 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      110 a    227 c    197 g    166 t
ORIGIN

  Query Match             12.7%;  Score 433.8;  DB 12;  Length 700;
  Best Local Similarity   84.5%;  Pred. No. 3.9e-48;
  Matches  582;  Conservative    0;  Mismatches   97;  Indels   10;  Gaps    8;

Qy        826 TGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGG 885
              ||||||||||||||| |  |||||||||||||||||| || ||||| || |||||| ||||
Db         16 TGACTGGGACACCAGCGTTCTGGCCCCCTACCTGGGTACTCAGGAAGAATGCCTC-TTGG 74

Qy        886 CCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGC 945
              ||| ||||||||||| ||||||| |||| | |||||||| ||| | ||| | ||||||||
Db         75 CCTCCTCACCCTCATTTTCCTCATCTGCATGGCAGCCACTCTG-TTGTGACGGAGGAGGC 133

Qy        946 AGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTG 1005
              ||  |||||||| ||||||| ||||||||| ||| || | |  || |||||| ||  ||
Db        134 AGTACTGGGCCCACCCGAGCCGGCAGAAGGGTTGT-GGTCTCTGCCGTGTCGCGCCGATG 192

Qy       1006 CTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCA 1065
              ||| |||||||  |   || |||||||||||| ||||| ||| |||||  ||||||||||||
Db        193 CTGCCCATGCCACGTTGGCCTGGCTTTCCGGAATCTGGGTACCCTGTTTCCCCGGCTGCA 252

Qy       1066 CCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAG 1125
               ||||||||||||||||||||| |||||||| ||||| ||||| ||||||||||||||||
Db        253 GCAGCTGTGCTGCCGCATGCCTCGCACCCTACGCCGACTCTTTGTGGCTGAGCTGTGCAG 312

Qy       1126 CTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTA 1185
              ||||||||||| ||||| ||||| ||||| |||||||| ||||||||| |||||||||||
Db        313 CTGGATGGCACTTATGACTTTCACACTGTTCTACACGGACTTCGTGGGAGAGGGGCTGTA 372

Qy       1186 CCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGT 1245
              |||| | || |||||||| ||||| |||||||||||||||||||||||||||||||||| |
Db        373 CCAGCGTGTACCCAGAGCCGAGCCAGGCACCGAGGCCCGGAGACACTATGATGAAGGCAT 432

Qy       1246 TCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGT 1305
              ||| ||||||||||||||||| ||||||||||| ||||||||||||||||||||  |||||
Db        433 TCGAATGGGCAGCCTGGGGCTCTTCCTGCAGTGTGCCATCTCCCTGGTCTTCTCCCTGGT 492

Qy       1306 CATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCA-GTGTGGCAG 1364
              |||||| ||||||| |||   |||||||| || |||||||| |||||| |||||
Db        493 CATGGACAGGCTGGTACAGAAGTTCGGCACACGGTCAGTCTATCTGGCCACGTGTGATGA 552

Qy       1365 CTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTT 1424
              | || ||||||||||||| |||||| || |||||||||| ||||  || |||||||| |
Db        553 CCTTTCCTGTGGCTGCCGTTGCCACCTG-CTGTCCCACAGCGTGGTGGTAGTGACAGCCT 611

Qy       1425 CAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCT 1484
              |||| | |||||||||||||||||||| | | ||||||||||||       |||    ||
Db        612 CAGCTG-CCTCACCGGGTTCACCTTCTC-GGCTTGCAGATCCTGC---TTACAGCTCGCT 666

Qy       1485 CCCTCTACCACCGGGAGAAGCAGGTGTTC 1513
              ||||||||||||| ||||||| ||||| |
```

```
Qy       3276 TCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTT 3324
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db         49 TCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTT 1

RESULT 29
BG081576
LOCUS       BG081576                 549 bp    mRNA    linear   EST 26-JAN-2001
DEFINITION  H3066G04-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
            H3066G04 5', mRNA sequence.
ACCESSION   BG081576
VERSION     BG081576.1  GI:12564144
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 549)
  AUTHORS   Kargul,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Jaradat,S.A., Tanaka
            ,T.S., Carter,M.G. and Ko,M.S.H.
  TITLE     Verification and initial annotation of NIA mouse 15K cDNA clone set
  JOURNAL   Unpublished (2001)
COMMENT     Other_ESTs: H3066G04-3
            Contact: George J. Kargul
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            This clone set has been freely distributed to the community. Please
            visit http://lgsun.grc.nia.nih.gov/cDNA/15k.html for details.
            Plate: H3066  row: G  column: 04
            Seq primer: -21M13 Reverse
            High quality sequence stop: 549
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1..549
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="niaEST:H3066G04-5"
                     /db_xref="taxon:10090"
                     /clone="H3066G04"
                     /clone_lib="NIA Mouse 15K cDNA Clone Set"
                     /sex="Clones arrayed from a variety of cDNA libraries"
                     /dev_stage="Clones arrayed from a variety of cDNA
                     libraries"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
                     clone is among a rearrayed set of 15,247 clones from 11
                     embryo  cDNA libraries (including preimplantation stage
                     embryos from  unfertilized egg to blastocyst, embryonic
                     part of E7.5 embryos,  extraembryonic part of E7.5 embryos
                     , and E12.5 female mesonephros/gonad)  and one newborn
                     ovary cDNA library. Average insert size 1.5 kb. All
                     source libraries are cloned unidirectionally with Oligo(dT
                     )-Not primers.  References include: (1) Genome-wide
                     expression profiling of mid-gestation  placenta and embryo
                     using a 15,000 mouse developmental cDNA microarray,  2000,
                     Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2)
                     Large-scale cDNA  analysis reveals phased gene expression
                     patterns during preimplantation  mouse develolpment, 2000,
                     Development, 127: 1737-1749; (3) Genome-wide  mapping of
                     unselected transcripts from extraembryonic tissue of
                     7.5-day  mouse embryos reveals enrichment in the t-complex
                     and under-representation  on the X chromosome, 1998, Hum
                     Mol Genet 7: 1967-1978."
BASE COUNT       98 a    176 c    160 g    115 t
ORIGIN

  Query Match             12.9%;  Score 438.6;  DB 12;  Length 549;
  Best Local Similarity   87.4%;  Pred. No. 1e-48;
  Matches  480;  Conservative    0;  Mismatches   69;  Indels    0;  Gaps    0;

Qy       1122 GCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGC 1181
              ||||||||||||||| ||||| ||||| ||||| |||||||| |||||||| |||||||
Db          1 GCAGCTGGATGGCACTTATGACTTTCACACTGTTCTACACGGACTTCGTGGGAGAGGGGC 60

Qy       1182 TGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAG 1241
              |||||||||| || |||||||| ||||| ||||||||||||||||||||||||||||||
Db         61 TGTACCAGGGTGTACCCAGAGCCGAGCCAGGCACCGAGGCCCGGAGACACTATGATGAAG 120

Qy       1242 GCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTC 1301
              || |||| |||||||||||||||| |||||||||| |||||||||||||||||||| |
Db        121 GCATTCGAATGGGCAGCCTGGGGCTCTTCCTGCAGTGTGCCATCTCCCTGGTCTTCTCCC 180

Qy       1302 TGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGG 1361
              ||||||||||| ||||||| |||    |||||||| ||  |||||||| ||||||||||
Db        181 TGGTCATGGACAGGCTGGTACAGAAGTTCGGCACACGGTCAGTCTATCTGGCCAGTGTGA 240

Qy       1362 CAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAG 1421
                 | || ||||||||||||| |||||| ||||||||||||| ||||   || |||||||
Db        241 TGACCTTTCCTGTGGCTGCCGCTGCCACCTGCCTGTCCCACAGCGTGGTGGTAGTGACAG 300

Qy       1422 CTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGG 1481
              | ||||| ||||||||||||||||||||||| ||| ||||||||||||| ||||| || |
Db        301 CCTCAGCTGCCCTCACCGGGTTCACCTTCTCGGCCTTGCAGATCCTGCCTTACACGCTCG 360

Qy       1482 CCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAG 1541
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||| |||||
Db        361 CCTCCCTCTACCACCGTGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACGCTGGAG 420

Qy       1542 GTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTC 1601
              ||    ||| ||||||||||||| ||  ||||||||| ||||||||||||||| ||||||
Db        421 GTAGCAGCGGTGAGGACAGCCAGACAACCAGCTTCTTGCCAGGCCCTAAGCCAGGAGCTC 480

Qy       1602 CCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGC 1661
                ||||| |||||||||||||||   ||||  |||| ||| | ||  | |||||||| || |
Db        481 TCTTCCCCAATGGACACGTGGGCTCTGGCAGCAGCGGCATCCTGGCCCCTCCACCTGCAC 540

Qy       1662 TCTGCGGGG 1670
              |||| ||||
Db        541 TCTGTGGGG 549

RESULT 30
BE674096
LOCUS       BE674096                 448 bp    mRNA    linear   EST 08-SEP-2000
DEFINITION  7d74g04.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3278742 3'
            similar to contains element MER22 repetitive element ;, mRNA
            sequence.
ACCESSION   BE674096
VERSION     BE674096.1  GI:10034637
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 448)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1..448
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
```

```
BASE COUNT      164 a    278 c    273 g    219 t
ORIGIN

  Query Match             13.3%;  Score 453;  DB 14;  Length 934;
  Best Local Similarity   77.5%;  Pred. No. 1e-50;
  Matches  636;  Conservative    0;  Mismatches  130;  Indels   55;  Gaps    5;

Qy         53 GGATCTGAGTGATGAGACGTGTCCCCACTGAG--GTGCCCCACAGCAGCAGGTGTTGAGC 110
Db         22 GGATCTGACCGACGAGATGTGTCCCCATCAAGCAAGGCACTAGATGGTGACGTGTTTAGC 81

Qy        111 ATGGGCTGAGAAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTC 170
Db         82 GTGGGACGAGATGCTGAATTGGCACTAAAGGGCTGGCAGAAATGGGAACCTGGCTGCACC 141

Qy        171 CTAGGCAGTTGGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGA 230
Db        142 CTAGGAGGTTAGTGCTAGTGAGGAGGAGAAGCCAC---------GGCAGGGCTGACTCAA 192

Qy        231 AGCAGTTCTGGAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCC 290
Db        193 AGCAGCTGTGGAGTATGTGAGTAGCCCCTGGAACCCTACCTGCCCTGTCCATCATGATCC 252

Qy        291 AGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACC 350
Db        253 AGAGGCTGTGGGCCAGCCGTCTGCTACGGCACCGGAAAGCTCAGCTCCTGCTGGTCAACC 312

Qy        351 TGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGC 410
Db        313 TGCTCACCTTTGGCCTGGAGGTGTGCCTGGCTGCCGGCATTACCTATGTGCCACCCCTTC 372

Qy        411 TGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGC 470
Db        373 TGCTGGAAGTCGGGGTGGAGGAGAAATTCATGACCATGGTGTTG---------------- 416

Qy        471 TGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATG 530
Db        417 -------------------------GGCTCAGCCAGTGACCAGTGGCGTGGGCGCTATG 450

Qy        531 GCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCA 590
Db        451 GCCGCCGGAGACCCTTTATCTGGGCTTTGTCCCTGGGTGTCCTGCTAAGCCTCTTTCTCA 510

Qy        591 TCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGG 650
Db        511 TCCCGAGGGCTGGCTGGCTGGCAGGACTGCTGTACCCAGACACCAGGCCCCTGGAGTTGG 570

Qy        651 CACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCAC 710
Db        571 CCCTGCTGATCTTGGGAGTGGGGCTGCTGGACTTTTGTGGCCAGGTGTGCTTTACTCCAT 630

Qy        711 TGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTG 770
Db        631 TGGAGGCCTTACTCTCCGACCTCTTCCGGGACCCAGACCACTGCCGCCAAGCCTTCTCTG 690

Qy        771 TCTATGCCTTCATGATCAGTC-TTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC 829
Db        691 TCTACGCCTTCATGATCAGCCTTTGGGGGCTGCCTGGGCTACCTCTTTACTGCCATTGAC 750

Qy        830 TGGGACACCAGTGCCCT-GGCCCCCTACCTGGGCACCCAGG 869
Db        751 TGGGACACCAGCGTTCTGGGCCCCCTACCTGGGTACTCAGG 791
```

```
RESULT 28
AI703348/c
LOCUS       AI703348                 469 bp    mRNA    linear   EST 18-DEC-1999
DEFINITION  wd93b09.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:2339129 3',
            mRNA sequence.
ACCESSION   AI703348
VERSION     AI703348.1  GI:4991248
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 469)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 566   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 447.
FEATURES             Location/Qualifiers
     source          1..469
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2339129"
                     /clone_lib="NCI_CGAP_Lu24"
                     /tissue_type="carcinoid"
                     /lab_host="DH10B"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Plasmid DNA from the normalized
                     library NCI_CGAP_Lu5 was prepared, and ss circles were
                     made in vitro. Following HAP purification, this DNA was
                     used as tracer in a subtractive hybridization reaction.
                     The driver was PCR-amplified cDNAs from a pool of 5,000
                     clones made from the same library (cloneIDs
                     1414920-1417991 and 1520904-1522439). Subtraction by Bento
                     Soares and M. Fatima Bonaldo. "
BASE COUNT      126 a     97 c    135 g    107 t      4 others
ORIGIN

  Query Match             13.2%;  Score 449.2;  DB 9;  Length 469;
  Best Local Similarity   98.5%;  Pred. No. 4.5e-50;
  Matches  462;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps    1;

Qy       2857 AAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGC 2916
Db        469 AAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGC 410

Qy       2917 CCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAG 2976
Db        409 CCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTATCAGGGACTGGNCTNATGAAG 350

Qy       2977 GCACTGCCCAAAATTT-CCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAG 3035
Db        349 GCACTGCNCAAAATTTNCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAG 290

Qy       3036 CTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGC 3095
Db        289 CTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGC 230

Qy       3096 CTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCA 3155
Db        229 CTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCA 170

Qy       3156 GGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGT 3215
Db        169 GGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGT 110

Qy       3216 TTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAA 3275
Db        109 TTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAA 50
```

```
Qy       1426 AGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTC 1485
Db        622 CAGTGCCCTCACGGGGTGCACCTTCTC-GGGCTGGAGATACCGGCTTACACGCACG-CTT 679

Qy       1486 CCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGC 1545
Db        680 CCTCTACCACCGTGCGAAGCAGTGTGTC--TGCCCAATACCGAGGGGCACCTGAGGTGAC 737

Qy       1546 TAGCAGTGAGGACAG 1560
Db        738 AAGAGATAGAGGCAG 752


RESULT 26
AW412402
LOCUS       AW412402              592 bp    mRNA    linear   EST 08-FEB-2000
DEFINITION  uo78h02.y1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:2648691 5'
            similar to TR:Q39231 Q39231 SUCROSE-PROTON SYMPORTER. [1] ;, mRNA
            sequence.
ACCESSION   AW412402
VERSION     AW412402.1  GI:6938274
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 592)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Other_ESTs: uo78h02.x1
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            MGI:1029143
            Seq primer: -40RP from Gibco
            High quality sequence stop: 421.
FEATURES             Location/Qualifiers
     source          1..592
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:2648691"
                     /clone_lib="NCI_CGAP_Mam1"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="10 months, virgin"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
BASE COUNT       91 a    193 c    168 g    140 t
ORIGIN

  Query Match            13.3%;  Score 455.2;  DB 10;  Length 592;
  Best Local Similarity   86.7%;  Pred. No. 6.5e-51;
  Matches  513;  Conservative    0;  Mismatches   78;  Indels    1;  Gaps    1;

Qy        728 GACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATC 787
Db          2 GACCTCTTCCGGGACCCAGACCACTGCCGCCAAGCCTTCTCTGTCTACGCCTTCATGATC 61

Qy        788 AGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTG 847
Db         62 AGCCTTGGGGGCTGCCTGGGCTACCTCTTACCTGCCATTGACTGGGACACCAGCGTTCTG 121

Qy        848 GCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTC 907
Db        122 GCCCCCTACCTGGGTACTCAGGAAGAATGCCTCTTTGGCCTCCTCACCCTCATTTTCCTC 181

Qy        908 ACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCA 967
Db        182 ATCTGCATGGCAGCCACTCTGTTTGTGACGGAGGAGGCAGTACTGGGCCCACCCGAGCCG 241

Qy        968 GCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTG 1027
Db        242 GCAGAAGGGTTGTTGGTCTCTGCCGTGTCGCGCCGATGCTGCCCATGCCACGTTGGCCTG 301

Qy       1028 GCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCC 1087
Db        302 GCTTTCCGGAATCTGGGTACCCTGTTTCCCCGGCTGCAGCAGCTGTGCTGCCGCATGCCT 361

Qy       1088 CGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTC 1147
Db        362 CGCACCCTACGCCGACTCTTTGTGGCTGAGCTGTGCAGCTGGATGGCACTTATGACTTTC 421

Qy       1148 ACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAG 1207
Db        422 ACACTGTTCTACACGGACTTCGTGGGAGAGGGGCTGTACCAGGGTGTACCCAGAGCCGAG 481

Qy       1208 CCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTG 1267
Db        482 CCAGGCACCGAGGCCCGGAGACACTATGATGAAGGCATTCGAATGGGGCAGCCTGGGCTC 541

Qy       1268 TTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGG 1319
Db        542 TTCCTGCAGTGTGCCATCTCCCTGGTCTTCTCCCT-GTCATGGACAGGCTGG 592


RESULT 27
BQ934815
LOCUS       BQ934815              934 bp    mRNA    linear   EST 21-AUG-2002
DEFINITION  AGENCOURT_8881033 NCI_CGAP_Co24 Mus musculus cDNA clone
            IMAGE:6396364 5', mRNA sequence.
ACCESSION   BQ934815
VERSION     BQ934815.1  GI:22350198
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 934)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13892  row: i  column: 05
            High quality sequence stop: 634.
FEATURES             Location/Qualifiers
     source          1..934
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6396364"
                     /clone_lib="NCI_CGAP_Co24"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.6 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
```

```
     source          1. .491
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1158206"
                     /clone_lib="NCI_CGAP_Pr22"
                     /sex="male"
                     /tissue_type="normal prostate"
                     /lab_host="DH10B"
                     /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; 1st strand cDNA was prepared
                     from normal prostate bulk tissue, and was then primed with
                     a Not I - oligo(dT) primer. Double-stranded cDNA was
                     ligated to Eco RI adaptors (Pharmacia), digested with Not
                     I and cloned into the Not I  and Eco RI sites of the
                     modified pT7T3 vector.  Library is normalized, and was
                     constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      135 a    104 c    136 g    116 t
ORIGIN

  Query Match            13.4%;  Score 456.8;  DB 9;  Length 491;
  Best Local Similarity   98.0%;  Pred. No. 4.4e-51;
  Matches  484;  Conservative    0;  Mismatches    7;  Indels    3;  Gaps    2;

Qy       2837 CAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTA 2896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        491 CAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTA 432

Qy       2897 ACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCT 2956
              |||||||||||| |||||  ||   |   |||||||| ||||||||||||||||||||||
Db        431 ACCACCCCTCTTTTCTTG--CCCAGCCTGTTCCCCCTACTTCCACTCCCCTCTACTCTCT 374

Qy       2957 CTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTAC 3016
              |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        373 CTAGGACTGGGCTGATGAA-GCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTAC 315

Qy       3017 CCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCAC 3076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 CCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCAC 255

Qy       3077 AAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGG 3136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 AAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGG 195

Qy       3137 AATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCA 3196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 AATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCA 135

Qy       3197 GGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATT 3256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATT 75

Qy       3257 TTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCT 3316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 TTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCT 15

Qy       3317 TATATGTTTAAAAA 3330
              ||||||||||||||
Db         14 TATATGTTTAAAAA 1


RESULT 25
BF785813
LOCUS       BF785813                 929 bp    mRNA    linear   EST 12-JAN-2001
DEFINITION  602112437F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4240617
            5', mRNA sequence.
ACCESSION   BF785813
VERSION     BF785813.1  GI:12090849
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 929)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9855  row: j  column: 10
            High quality sequence stop: 610.
FEATURES             Location/Qualifiers
     source          1. .929
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4240617"
                     /clone_lib="NCI_CGAP_Kid14"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library. |"
BASE COUNT      190 a    275 c    262 g    202 t
ORIGIN

  Query Match            13.4%;  Score 456.2;  DB 12;  Length 929;
  Best Local Similarity   80.8%;  Pred. No. 3.9e-51;
  Matches  594;  Conservative    0;  Mismatches  133;  Indels    8;  Gaps    5;

Qy        826 TGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGG 885
              ||||||||||||||  |   ||||||||||||||||| || |||||  || |||||||| ||
Db         26 TGACTGGGACACCAGCGTTCTGGCCCCCTACCTGGGTACTCAGGAAGAATGCCTCTTAGG 85

Qy        886 CCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGC 945
              ||| ||||||||||| ||||||| |||| | |||||||| ||| | ||| | ||||||||
Db         86 CCTCCTCACCCTCATTTTCCTCATCTGCATGGCAGCCACTCTGTTTGTGACGGAGGAGGC 145

Qy        946 AGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTG 1005
              ||  ||||||||  |||||| ||||||||| || | |  | ||  || |||||| || ||
Db        146 AGTACTGGGCCCACCCGAGCCGGCAGAAGGGTTGTTGGTCTCTGCCGTGTCGCGCCGATG 205

Qy       1006 CTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCA 1065
              ||| |||||||   |     || ||||||||||||| |||||   ||||| ||||||||||||
Db        206 CTGCCCATGCCACGTTGGCCTGGCTTTCCGGAATCTGGGTACCCTGTTTCCCCGGCTGCA 265

Qy       1066 CCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAG 1125
               |||||||||||||||||| |||||||| ||||| ||||| ||||||||||||||||||
Db        266 GCAGCTGTGCTGCCGCATGCCTCGCACCCTACGCCGACTCTTTGTGGCTGAGCTGTGCAG 325

Qy       1126 CTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTA 1185
              ||||||||||| ||||| ||||| ||||| |||||||| |||||||| ||||||||||
Db        326 CTGGATGGCACTTATGACTTTCACACTGTTCTACACGGACTTCGTGGGAGAGGGGCTGTA 385

Qy       1186 CCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGT 1245
              |||||| || |||||||| ||||||||||||||||||||||||||||||||||||||| |
Db        386 CCAGGGTGTACCCAGAGCCGAGCCAGGCACCGAGGCCCGGAGACACTATGATGAAGGCAT 445

Qy       1246 TCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGT 1305
              ||| |||||||||||||||| |||||||||| |||||||||||||||||||| |||||
Db        446 TCGAATGGGCAGCCTGGGGCTCTTCCTGCAGTGTGCCATCTCCCTGGTCTTCTCCCTGGT 505

Qy       1306 CATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGC 1365
              ||||||| |||||| |||    |||||||| ||  ||||||||| ||||||||||    |
Db        506 CATGGACAGGCTGGTACAGAAGTTCGGCACACGGTCAGTCTATCTGGCCAGTGTGATGAC 565

Qy       1366 TTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTC 1425
               || |||||||||||  |       || |||||||||| ||||  || ||||||||  |
Db        566 CTTTCCTGTGGCTGC--GCTGTCACTGTCTGTCCCACAGCGTGGTGGTAGTGACAGC--C 621
```

```
Qy          965 CCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGC 1024
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          302 CCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGC 361

Qy         1025 TTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATG 1084
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          362 TTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATG 421

Qy         1085 CCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACC 1144
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          422 CCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACC 481

Qy         1145 T 1145
                |
Db          482 T 482
```

RESULT 23
BG173136

```
LOCUS       BG173136                 969 bp    mRNA    linear   EST 06-FEB-2001
DEFINITION  602335411F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4458602 5',
            mRNA sequence.
ACCESSION   BG173136
VERSION     BG173136.1  GI:12679748
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 969)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10257  row: e  column: 03
            High quality sequence stop: 608.
FEATURES             Location/Qualifiers
     source          1..969
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4458602"
                     /clone_lib="NCI_CGAP_Mam1"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="10 months, virgin"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
BASE COUNT      199 a    289 c    287 g    194 t
ORIGIN

  Query Match             13.5%;  Score 459;  DB 12;  Length 969;
  Best Local Similarity   82.9%;  Pred. No. 1.6e-51;
  Matches  583;  Conservative    0;  Mismatches  115;  Indels    5;  Gaps    5;

Qy          839 AGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTC 898
                 || |  ||||||||||||||||| || ||||| || |||||||||||||| ||||||||
Db            1 AGCGTTCTGGCCCCCTACCTGGGTACTCAGGAAGAATGCCTCTTTGGCCTCCTCACCCTC 60

Qy          899 ATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCC 958
                || ||||||| |||| | |||||||| ||| | ||| ||||||||||  ||||||||
Db           61 ATTTTCCTCATCTGCATGGCAGCCACTCTGTTTGTGACGGAGGAGGCAGTACTGGGCCCA 120

Qy          959 ACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGG 1018
                 |||||| |||||||| ||| || |  | || |||||| ||  | ||||| |||||||
Db          121 CCCGAGCCGGCAGAAGGGTTGTTGGTCTCTGCCGTGTCGCGCCGATGCTGCCCATGCCAC 180

Qy         1019 GCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGC 1078
                |    || |||||||||||||| ||||| ||||| ||||||||||||| ||||||||||||
Db          181 GTTGGCCTGGCTTTCCGGAATCTGGGTACCCTGTTTCCCCGGCTGCAGCAGCTGTGCTGC 240

Qy         1079 CGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTC 1138
                ||||||| |||||||| ||||| ||||| |||||||||||||||||||||||||||||
Db          241 CGCATGCCTCGCACCCTACGCCGACTCTTTGTGGCTGAGCTGTGCAGCTGGATGGCACTT 300

Qy         1139 ATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCC 1198
                ||||| ||||| ||||| |||||||| |||||||| ||||||||||||| || || |||
Db          301 ATGACTTTCACACTGTTCTACACGGACTTCGTGGGAGAGGGGCTGTACCA-GGTGTACCC 359

Qy         1199 AGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGC 1258
                ||||| ||||| |||||||||| ||||||||||||||||||||||| |||| |||||||||
Db          360 AGAGCCGAGCCAGGCACCGAGG-CCGGAGACACTATGATGAAGGCATTCGAATGGGCAGC 418

Qy         1259 CTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTG 1318
                ||||||| ||||||||||| |||||||||||||||||||| |||||||||||| |||||
Db          419 CTGGGGCTCTTCCTGCAGTGTGCCATCTCCCTGGTCTTCTCCCTGGTCATGGACAGGCTG 478

Qy         1319 GTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCT 1378
                || |||   |||||||| ||  |||||||| |||||||||||      | || ||||||||
Db          479 GTACAGAAGTTCGGCACACGGTCAGTCTATCTGGCCAGTGTGATGACCTTTCCTGTGGCT 538

Qy         1379 GCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACC 1438
                |||| |||||| || ||||||||||| ||||   || |||||||| ||||   ||||||||
Db          539 GCCGCTGCCACCTG-CTGTCCCACAGCGTGGTGGTAGTGACAGC-TCAGGTGCCCTCACC 596

Qy         1439 GGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGG 1498
                ||||||||||||| | | ||||||||||| || | | | |||||| |||||  |  |
Db          597 GGGTTCACCTTCTCCGGCTTGCAGATCCTGACTTAAAGA-TCGCCTCCTTCTACACCGTG 655

Qy         1499 GAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAG 1541
                  ||||   |  | | | ||||||| ||  ||||| ||| ||
Db          656 AGAAAGCCACTTGTTCCGGCCAAATAACGGAGGGACCCTGAAG 698
```

RESULT 24
AA631143/c

```
LOCUS       AA631143                 491 bp    mRNA    linear   EST 31-OCT-1997
DEFINITION  nq75g08.s1 NCI_CGAP_Pr22 Homo sapiens cDNA clone IMAGE:1158206 3',
            mRNA sequence.
ACCESSION   AA631143
VERSION     AA631143.1  GI:2553754
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 491)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 919   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 472.
FEATURES             Location/Qualifiers
```

```
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11855  row: p  column: 02
            High quality sequence stop: 778.
FEATURES             Location/Qualifiers
     source          1..901
                     /organism="Mus musculus"
                     /strain="129,C57BL/6J,FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5337073"
                     /clone_lib="NCI_CGAP_Mam3"
                     /tissue_type="tumor, gross tissue"
                     /dev_stage="10 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
                     Reference for transgenic model: Xu et al., Nature Genetics
                     22, 37-43 (1999)."
BASE COUNT      151 a    260 c    288 g    202 t
ORIGIN

  Query Match             13.6%;  Score 463;  DB 13;  Length 901;
  Best Local Similarity   78.5%;  Pred. No. 5e-52;
  Matches  684;  Conservative    0;  Mismatches  165;  Indels   22;  Gaps   10;

Qy          1 GGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGA 60
                | |  | |||  ||||||||  || ||  |||||||||||| ||||  |||||||||||||
Db         42 GTGTCCGAGCTCGCACGCGCCAGCCCCAGGTGACAGCCGCACGCC-GGGCCAGGATCTGA 100

Qy         61 GTGATGAGACGTGTCCCCACTGAGGT---GCCCCACAGCAGCAGGTGTTGAGCATGGGCT 117
                 || ||||  |||||||||  | | |       || | | |       | |||||  ||| ||||
Db        101 CCGACGAGATGTGTCCCCAGTCAAGCAAGGCACTAGATGGTGACGTGTTTAGCGTGGGAC 160

Qy        118 GAGAAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCA 177
              |||| ||||  |     |||||  |||||||||||||||||||  ||||||||     ||||||
Db        161 GAGATGCTGAATTGGCACTAAAGGGCTGGCAGAAATGGGAACCTGGCTGCACCCTAGGAG 220

Qy        178 GTTGGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTT 237
               ||| |  |   ||  |||||||||  ||| |              |||| || ||  | |||||| |
Db        221 GTTAGTGCTAGTGAGGAGGAGAAGCCAC---------GGCAGGGCTGACTCAAAGCAGCT 271

Qy        238 CTGGAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCT 297
               ||||||    |||    |||||   || ||||||| |||   |  |||    ||| ||||||||||
Db        272 GTGGAGTATGTGAGTAGCCCCTGGAACCCTACCTGCCCTGTCCATCATGATCCAGAGGCT 331

Qy        298 GTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAAC 357
              |||||    |||||  |||||||||||||||| ||||||  |||||||||||||||  ||
Db        332 GTGGGCCAGCCGTCTGCTACGGCACCGGAAAGCTCAGCTCCTGCTGGTCAACCTGCTCAC 391

Qy        358 CTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGA 417
              ||||||||||||||||    ||||  |||||  ||||||||||  || ||  ||||||||
Db        392 CTTTGGCCTGGAGGTGTGCCTGGCTGCCGGCATTACCTATGTGCCACCCCTTCTGCTGGA 451

Qy        418 AGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCT 477
              ||| ||||||  ||||||||  |||||||||||||||||  |||||||||||  ||||||||  |||||
Db        452 AGTCGGGGTGGAGGAGAAATTCATGACCATGGTGTTGGGCATTGGCCCAGTGCTAGGCCT 511

Qy        478 GGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCG 537
              |||  |  |||  ||  |||||||||||||||||||||||  ||||||||  |||||||||||||||
Db        512 GGTTTCTGTTCCACTCCTAGGCTCAGCCAGTGACCAGTGGCGTGGGCGCTATGGCCGCCG 571

Qy        538 CCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCCTCTTTCTCATCCCAAG 597
                 | |||||  |||||||  |||||  ||||  ||||||||||||||||||||||||||  ||
Db        572 GAGACCCTTTATCTGGGCTTTGTCCCTGGGTGTCCTGCTAAGCCCTCTTTCTCATCCCGAG 631

Qy        598 GGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCT 657
              |||  |||||||||  |||||  |||||||  |||  ||        |||||||||||||||  ||||
Db        632 GGCTGGCTGGCTGGCAGGACTGCTGTACCCAGACATCAGGCCCCTGGAGTTGGCCCTGCT 691

Qy        658 CATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCT--TCACTCCACTGGAG 715
                |||  ||||  |||||||||||||||  |  ||||||||||||||||     |      |     |||||
Db        692 GATCTTGGGAGTGGGGCTGCTGGACGT-TGTGGCCAGGTGTGCTTGTACTCCATTTGGAG 750

Qy        716 GCCCTGCTCTCTGA--CCTCTTCCGGGACCCGGACCAC-TGTCGCCAGGCCTACTCTGTC 772
              ||| | |||||  ||    |       | |  ||||||| ||||||  ||  |||  ||||||  |||||||
Db        751 GCCTTACTCTCCGAACCTCTTCCGGGACCCAGACCACTTGCCGCAAGGCCTTCTCTGTC 810

Qy        773 -TATGCCTTCATGATCAGTCTTGGGGGCTGCCT-GGGCTACCTCCTGCCTGCCATTGACT 830
                || ||||||  |||||||  |||||||||||||||   |||  |||||  ||    ||||      ||   |
Db        811 TTACGCCTTCCTGATCAGCCTTGGGGGCTGCCTGGGGTTACCTTCTTACTGCGCGTGCTT 870

Qy        831 GGGACA-CCAGTGCCCTGGCCCCCTACCTGG 860
              |||||| ||||  |    |||||||||||    ||
Db        871 GGGACACCCAGCGTTCTGGCCCCCTACTGGG 901
```

```
RESULT 22
AF109299
LOCUS       AF109299               482 bp    mRNA    linear   EST 28-JAN-2000
DEFINITION  AF109299 Homo sapiens prostate adult Homo sapiens cDNA clone IPCA-2
            , mRNA sequence.
ACCESSION   AF109299
VERSION     AF109299.1  GI:6782692
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 482)
  AUTHORS   Walker,M.G., Volkmuth,W., Sprinzak,E., Hodgson,D. and Klingler,T.
  TITLE     Prediction of gene function by genome-scale expression analysis:
            prostate cancer-associated genes
  JOURNAL   Genome Res. 9 (12), 1198-1203 (1999)
  MEDLINE   20082966
COMMENT     Contact: Walker MG
            Incyte Pharmaceuticals
            3174 Porter Drive, Palo Alto, CA 94304, USA
            co-expressed with known prostate-cancer genes.
FEATURES             Location/Qualifiers
     source          1..482
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IPCA-2"
                     /clone_lib="Homo sapiens prostate adult"
                     /tissue_type="prostate"
                     /dev_stage="adult"
                     /note="multiple clone assembly from multiple libraries and
                     donors"
BASE COUNT       57 a    179 c    141 g    105 t
ORIGIN

  Query Match             13.6%;  Score 462.6;  DB 9;  Length 482;
  Best Local Similarity   99.0%;  Pred. No. 7.6e-52;
  Matches  476;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy        666 GCGTGGGGCTGCTGGACTTCTGT-GGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTC 724
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db          2 GCGTGGGGCTGCTGGACTTCTGTGGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTC 61

Qy        725 TCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATG 784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 TCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATG 121

Qy        785 ATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCC 844
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 ATTAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCC 181

Qy        845 CTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTC 904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 CTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTC 241

Qy        905 CTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAG 964
              ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CTCACCTGCGTAGAGCCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAG 301
```

```
Qy        546 TCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCT 605
              | ||||||||  ||||| ||||  ||||||| |||||||||||||||||| ||||| ||||
Db        542 TTATCTGGGCTTTGTCCCTGGGTGTCCTGCTAAGCCTCTTTCTCATCCCGAGGGCTGGCT 601

Qy        606 GGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGG 665
              |||| ||||| |||||||  || ||  ||||||||||||| |||| ||||| ||| |||
Db        602 GGCTGGCAGGACTGCTGT-ACCAGACACCAGGCCCCTGGAGTTGGCCCTGCTGATCTTGG 660

Qy        666 GCGTGGGGCTGCTGGACTTCTG--TGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCT 723
              | |||||||||||||||| ||     | | |||||||||| |||||| |||||||| | ||
Db        661 GAGTGGGGCTGCTGGACTTTTGGTGGCCAAGGTGTGCTTTACTCCATTGGAGGCCTTACT 720

Qy        724 CTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTC-TGTCTATGCCTTCA 782
              ||| || |||||||||||| |.|| || || ||| | |||| ||| |||||| |||||||
Db        721 CTCCGAGCTCTTCCGGGACGCAGAACATTGGCGCAAAGCCTTCTCTTGTCTACGCCTTCA 780

Qy        783 TGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTG 842
              |||||||  ||||||||||||||||.|||| |||  |   | |||| ||||||||||| |||| || |
Db        781 TGATCAGCCTTGGGGGCTGCCTGGGTTACTTTTTACCTGGCATTGACTGGAACACAAG-G 839

Qy        843 CCCTGGCCCCCTACCTGGGCACCCAGG 869
                 |||| |||||||||||| |   ||||
Db        840 TGCTGGGCCCCTACCTGGGTAATCAGG 866
```

```
RESULT 20
BG174399
LOCUS       BG174399                 786 bp    mRNA    linear   EST 06-FEB-2001
DEFINITION  602334219F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4457452 5',
            mRNA sequence.
ACCESSION   BG174399
VERSION     BG174399.1  GI:12681102
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 786)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10254  row: e  column: 05
            High quality sequence stop: 666.
FEATURES             Location/Qualifiers
     source          1..786
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4457452"
                     /clone_lib="NCI_CGAP_Mam1"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="10 months, virgin"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
BASE COUNT      121 a    246 c    234 g    185 t
ORIGIN

  Query Match            14.7%;  Score 499.6;  DB 12;  Length 786;
  Best Local Similarity   85.7%;  Pred. No. 8.1e-57;
  Matches  579;  Conservative    0;  Mismatches   94;  Indels    3;  Gaps    2;
```

```
Qy        593 CCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCA 652
              || ||||| ||||||||| ||||| ||||||| ||| ||  |||||||||||||  |||
Db          1 CCGAGGGCTGGCTGGCTGGCAGGACTGCTGTACCCAGAGACCAGGCCCCTGGAGTAGGCC 60

Qy        653 CTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTG 712
              ||||| ||| |||| |||||||||||||||| |||||||||||||||| |||||| ||
Db         61 CTGCTGATCTTGGGAGTGGGGCTGCTGGACTTTTGTGGCCAGGTGTGCTTTACTCCATTG 120

Qy        713 GAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCT-GT 771
              |||||| | ||||| |||||||||||||||| |||||||| ||||| |||| |||| |
Db        121 GAGGCCTTACTCTCCGACCTCTTCCGGGACCCAGACCACTGCCGCCAAGCCTTCTCTGGG 180

Qy        772 CTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTG 831
                || ||||||||||||| |||||||||||||||||||||||| | |||||||||||||
Db        181 GTACGCCTTCATGATCAGCCTTGGGGGCTGCCTGGGCTACCTCTTACCTGCCATTGACTG 240

Qy        832 GGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCT 891
              |||||||| |  |||||||||||||||||| || ||||| ||||||||||||||||| ||
Db        241 GGACACCAGCGTTCTGGCCCCCTACCTGGGTACTCAGGAAGAATGCCTCTTTGGCCTCCT 300

Qy        892 CACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCT 951
              |||||||| ||||||| |||| | |||||||| ||| | ||| | |||||||||| ||
Db        301 CACCCTCATTTTCCTCATCTGCATGGCAGCCACTCTGTTTGTGACGGAGGAGGCAGTACT 360

Qy        952 GGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCC 1011
              ||||||  ||||||| ||||||||| ||| || | |  || |||||| ||  ||||| ||
Db        361 GGGCCCACCCGAGCCGGCAGAAGGGTTGTTGGTCTCTGCCGTGTCGCGCCGATGCTGCCC 420

Qy       1012 ATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCT 1071
              |||||  |     || ||||||||||||| |||||  ||||| |||||||||||| |||||
Db        421 ATGCCACGTTGGCCTGGCTTTCCGGAATCTGGGTACCCTGTTTCCCCGGCTGCAGCAGCT 480

Qy       1072 GTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGAT 1131
              |||||||||||||||| ||||||| ||||| ||||| ||||||||||||||||||||||||
Db        481 GTGCTGCCGCATGCCTCGCACCCTACGCCGACTCTTTGTGGCTGAGCTGTGCAGCTGGAT 540

Qy       1132 GGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGG 1191
              |||||| ||||| ||||| ||||| |||||||| ||||||||| |||||||||||| |  ||
Db        541 GGCACTTATGACTTTCACACTGTTCTACACGGACTTCGTGGGAGAGGGGCTGTAACCAGG 600

Qy       1192 CGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGA- 1250
                || |||||||| ||||| |||||||||||||||||||||||||||||||||| |||| |
Db        601 TGTACCCAGAGCCGAGCCAGGCACCGAGGCCCGGAGACACTATGATGAAGGCATTCGAAA 660

Qy       1251 -TGGGCAGCCTGGGGC 1265
                ||||||||||||||
Db        661 TGGGGCAGCCTGGGGC 676
```

```
RESULT 21
BI650119
LOCUS       BI650119                 901 bp    mRNA    linear   EST 12-SEP-2001
DEFINITION  603296208F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5337073 5',
            mRNA sequence.
ACCESSION   BI650119
VERSION     BI650119.1  GI:15564355
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 901)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
```

```
                  /clone="IMAGE:5473696"
                  /clone_lib="NIH_MGC_41"
                  /tissue_type="amelanotic melanoma, cell line"
                  /lab_host="DH10B (phage-resistant)"
                  /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                  EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                  into EcoRI/XhoI sites using the following 5' adaptor:
                  GGCACGAG(G). Library constructed by Ling Hong in the
                  laboratory of Gerald M. Rubin (University of California,
                  Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                  Superscript II RT (Life Technologies). Note: this is a
                  NIH_MGC Library."
BASE COUNT      310 a    665 c    258 g    418 t     16 others
ORIGIN

  Query Match             15.0%;  Score 512.8;  DB 14;  Length 1667;
  Best Local Similarity   97.4%;  Pred. No. 1e-58;
  Matches  553;  Conservative    0;  Mismatches   12;  Indels    3;  Gaps    3;

Qy       1802 TCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCT 1861
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCT 60

Qy       1862 GCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGC 1921
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGC 120

Qy       1922 GACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTC 1981
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTC 180

Qy       1982 ACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTT 2041
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTT 240

Qy       2042 TCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAG 2101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAA 300

Qy       2102 CTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGA 2161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGA 360

Qy       2162 CTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCA 2221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCA 420

Qy       2222 TGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAG-C 2280
              |||||||||||| | ||||||||||||||||||||| |||||  ||||||||||||| |
Db        421 TGCACTGGAATGCCGCGACTCTGCAGGTGGATTACCCACGCTCACGGTTAACAGCTAGCC 480

Qy       2281 CTCCTAGTTGAGACACACCTA-GAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCT 2339
              ||||| |||||||||||||| |||||||| | ||||||||||||||||||  ||||||
Db        481 CTCCTACTTGAGACACACCTACCAGAAGGGCTCTTGGGAGCTGAATAAACTCCTTCACCT 540

Qy       2340 GG-TTTCCCATCTCTAAGCCCCTTAACC 2366
              || |||||||||||||||| ||||||
Db        541 GGTTTTCCCATCTCTAAGCCCTTTAACC 568


RESULT 19
BG864609
LOCUS       BG864609                 872 bp    mRNA    linear   EST 29-MAY-2001
DEFINITION  602798469F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4919513 5',
            mRNA sequence.
ACCESSION   BG864609
VERSION     BG864609.1  GI:14215147
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 872)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
            Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10834  row: i  column: 18
            High quality sequence stop: 738.
FEATURES             Location/Qualifiers
     source          1. .872
                     /organism="Mus musculus"
                     /strain="NMRI"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4919513"
                     /clone_lib="NCI_CGAP_Mam4"
                     /tissue_type="tumor, gross tissue"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Priscilla Furth,
                     NIH Reference for transgenic model: Li et al., Cell Growth
                     and  Differentiation 7, 3-11 (1996)."
BASE COUNT      157 a    242 c    276 g    197 t
ORIGIN

  Query Match             14.9%;  Score 508.6;  DB 12;  Length 872;
  Best Local Similarity   79.7%;  Pred. No. 5.1e-58;
  Matches  691;  Conservative    0;  Mismatches  159;  Indels   17;  Gaps    7;

Qy          8 AGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTGATGA 67
              |||  |||||||  || ||  |||||||||||| ||||  ||||||||||||  || ||
Db         12 AGCTCGCACGCGCCAGCCCCAGGTGACAGCCGCACGCC-GGGCCAGGATCTGACCGACGA 70

Qy         68 GACGTGTCCCCACTGAGGT--GCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAGCT 125
              || ||||||||   ||     || | | |     | ||||| ||| ||||  |||| |||
Db         71 GATGTGTCCCCATCAAGCAAGGCACTAGATGGTGACGTGTTTAGCGTGGGACGAGATGCT 130

Qy        126 GGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGCGG 185
              | |  ||||| |||||||||||||||||||  ||||||||   ||||||  ||| | |
Db        131 GAATTGGCACTAAAGGGCTGGCAGAAATGGGAACCTGGCTGCACCCTAGGAGGTTAGTGC 190

Qy        186 CAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAGTG 245
                ||  ||||||||| ||| |          |||| || ||  | |||||| | ||||||
Db        191 TAGTGAGGAGGAGAAGCCAC---------GGCAGGGCTGACTCAAAGCAGCTGTGGAGTA 241

Qy        246 CCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGTGA 305
              |||    ||||| || ||||||| |||   | |||  |||  |||||||||||||||  |
Db        242 TGTGAGTAGCCCCTGGAACCCTACCTGCCCTGTCCATCATGATCCAGAGGCTGTGGGCCA 301

Qy        306 GCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCC 365
              |||| ||||| ||||||||||||||  ||||| |||||||||||||||| || |||||||
Db        302 GCCGTCTGCTACGGCACCGGAAAGCTCAGCTCCTGCTGGTCAACCTGCTCACCTTTGGCC 361

Qy        366 TGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGG 425
              |||||||||  |||| || ||||| |||||||||||  | || ||||||||||| ||||
Db        362 TGGAGGTGTGCCTGGCTGCCGGCATTACCTATGTGCCAGCCCTTCTGCTGGAAGTCGGGG 421

Qy        426 TAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTG 485
              | ||||||| ||||||||||||||| |||||||||| |||||||| |||||||| | ||
Db        422 TGGAGGAGAAATTCATGACCATGGTGTTGGGCATTGGCCCAGTGCTAGGCCTGGTTTCTG 481

Qy        486 TCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCT 545
              | || |||||||||||||||||||||| |||||||| |||||||||||||  | ||||
Db        482 TTCCACTCCTAGGCTCAGCCAGTGACCAGTGGCGTGGGCGCTATGGCCGCCGGAGACCCT 541
```

```
Db        302 CTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCAT 361

Qy       2724 TGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACC 2783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACC 421

Qy       2784 CAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTT 2843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 CAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTT 481

Qy       2844 AGGGTGTTGAA-GGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCT-AACCAC 2901
              ||||||||||| ||||||||||||||||||||||||||||||||||||||||| ||||||
Db        482 AGGGTGTTGAACGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAAACCAC 541

Qy       2902 CCCTCTTCTCTTGGCCCAGCCT-GGTTCCCCCCACTTCCACTCCCCTC 2948
              |||||||||||||||||||||| ||||||||||| |||||||||||||
Db        542 CCCTCTTCTCTTGGCCCAGCCTGGGTTCCCCCCATTTCCACTCCCCTC 589


RESULT 17
BG122427
LOCUS       BG122427                 715 bp    mRNA    linear   EST 30-JAN-2001
DEFINITION  602353324F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4451572 5',
            mRNA sequence.
ACCESSION   BG122427
VERSION     BG122427.1  GI:12615936
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 715)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10238  row: p  column: 05
            High quality sequence stop: 689.
FEATURES             Location/Qualifiers
     source          1..715
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4451572"
                     /clone_lib="NIH_MGC_90"
                     /tissue_type="adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.7 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      176 a    174 c    175 g    190 t
ORIGIN

  Query Match            15.4%;  Score 525;  DB 12;  Length 715;
  Best Local Similarity   94.5%;  Pred. No. 3.9e-60;
  Matches  664;  Conservative    0;  Mismatches   25;  Indels   14;  Gaps   11;

Qy       2211 GAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTT 2270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTT 60

Qy       2271 AACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTGGGAGCTGAATAAACT 2330
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         61 AACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG-TTTTGGGAGCTGAATAAACT 119

Qy       2331 CAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCT 2390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 CAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCT 179

Qy       2391 TGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG-ACTTATTT 2449
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |  |  ||
Db        180 TGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGAAAGTTATT 239

Qy       2450 GTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACT 2509
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        240 GTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAG-CCACAGCACT 298

Qy       2510 GTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTC 2569
              ||| |||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        299 GTC-TTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTG-TGGTCCTTC 356

Qy       2570 TGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGA 2629
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 TGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGA 416

Qy       2630 AGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGG-TGGGGGAT 2688
              ||||||||||||||||||||| |||||| ||||||||||||| |||||||| ||||||||
Db        417 AGGGAATCCATTGCTAGCTTTACTGTGTGGGTGTCTAATATTAGGGTAGGGCTGGGGGAT 476

Qy       2689 CCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATC-TT 2747
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||| ||
Db        477 CCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGCGCTGATCATTGCCAGAATCTTT 536

Qy       2748 CTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCAT 2807
              ||||||||||||||||||||||  |||   ||||||||||| || |||||||||||||||
Db        537 CTTCTCCTGGGGTCTGGCCCCC--AAAATGCTAACCCAGGA-CTGGGAAATTCTACTCAT 593

Qy       2808 CCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGG 2867
               ||||||||||||||||||||||||||||  | | |||||||||||||||||||||||||
Db        594 GCCAAATGATAATTCCAAATGCTGTTACC--AAGGTAGGGTGTTGAAGGAAGGTAGAGGG 651

Qy       2868 TGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCT 2910
               |||  |   |||||||||||  | |||||||||||   ||||
Db        652 GGGGCTTCAGGGTCTCAACGG--TACCTAACCACCCTCTTTCT 692


RESULT 18
BM912193
LOCUS       BM912193                1667 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6613122 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5473696
            5', mRNA sequence.
ACCESSION   BM912193
VERSION     BM912193.1  GI:19362572
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1667)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1985  row: l  column: 17
            High quality sequence stop: 448.
FEATURES             Location/Qualifiers
     source          1..1667
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
```

```
          High quality sequence start: 3
          High quality sequence stop: 602.
FEATURES             Location/Qualifiers
     source          1..1035
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4905674"
                     /clone_lib="NIH_MGC_17"
                     /tissue_type="rhabdomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
                     Site_2: XhoI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G).  Size-selected >500bp
                     for average insert size 1.8kb.  Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      227 a    273 c    341 g    194 t
ORIGIN

  Query Match             15.8%;  Score 538.2;  DB 12;  Length 1035;
  Best Local Similarity   93.1%;  Pred. No. 5.9e-62;
  Matches  632;  Conservative    0;  Mismatches   33;  Indels   14;  Gaps    6;

Qy          7 CAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTGATG 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTGATG 61

Qy         67 AGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAGCTG 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 AGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAGCTG 121

Qy        127 GACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGCGGC 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGCGGC 181

Qy        187 AGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAGTGC 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 AGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAGTGC 241

Qy        247 CTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGTGAG 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGTGAG 301

Qy        307 CCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCT 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 CCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCT 361

Qy        367 GGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGT 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGT 421

Qy        427 AGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGT 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 AGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGT 481

Qy        487 CCCGC-TCCTAGGCTCA-GCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCC 544
              ||||| ||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        482 CCCGCATCCTAGGCTCACGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCC 541

Qy        545 TTCATCTGGGCACTGTCCTTGGGCATCCTGCTGA------GCCTCTTTCTCATCCCAAGG 598
              |||||||||        | |  ||   | |  ||||       |||||||||||||||
Db        542 TTCATCTGGAGCACTGTCCTTGGAGGCATCCTGACTTGAGCATTCTTTCTCATCCCAAGG 601

Qy        599 GCCGGCTGGCTAGCA----GGGCTGCTGTGCCC-GGATCCCAGGCCCCTGGAGCTGGCAC 653
              |||||    |      |     || ||||||||||| ||||  |||||  ||||||||||||| ||
Db        602 GCCGGGCTGGCTAGACAAGGGACTGCTGTGCCCAGGATACCAGGACCCTGGAGCTGG-AC 660

Qy        654 TGCTCATCCTGGGCGTGGG 672
              ||||||| |||||    ||
Db        661 TGCTCATCATGGGCAGTGG 679
```

```
RESULT 16
BG469586
LOCUS       BG469586               589 bp    mRNA    linear   EST 21-MAR-2001
DEFINITION  602533622F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4661500 5',
            mRNA sequence.
ACCESSION   BG469586
VERSION     BG469586.1  GI:13401861
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 589)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: NIH Intramural Sequencing Center
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1460  row: k  column: 05
            High quality sequence stop: 587.
FEATURES             Location/Qualifiers
     source          1..589
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4661500"
                     /clone_lib="NIH_MGC_15"
                     /tissue_type="adenocarcinoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)"
BASE COUNT      135 a    153 c    133 g    168 t
ORIGIN

  Query Match             15.7%;  Score 534.8;  DB 12;  Length 589;
  Best Local Similarity   98.8%;  Pred. No. 2.2e-61;
  Matches  581;  Conservative    0;  Mismatches    2;  Indels    5;  Gaps    4;

Qy       2366 CTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCAT 2425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCAT 61

Qy       2426 GGGATTTGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGA 2483
              ||||||||||||||||  | ||||||||||||||||||||||||||||||||||||||||
Db         62 GGGATTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGA 121

Qy       2484 ACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTT 2543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 ACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTT 181

Qy       2544 ATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATAT 2603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 ATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATAT 241

Qy       2604 TTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGT 2663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 TTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGT 301

Qy       2664 CTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCAT 2723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        301 AGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTG 360

Qy       1896 CTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCA 1955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCA 420

Qy       1956 GCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCA 2015
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        421 GCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTACCCCCA 480

Qy       2016 TGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTG 2075
              |||||||||||| ||||||| ||| ||||||||||||||||||  |||||||||||| |||
Db        481 TGGGGCTGCCGGCCTGGCCGGCAGGTTCTGTTGCTGCCAAAGAAATGTGGCTCTCTACTG 540

Qy       2076 CCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGG 2115
              |||||||||||||||||| ||  ||||||||| | ||||
Db        541 CCACCCTGTGCTGCTGAGGAGCCTAGCTGCACAACGGGGG 580


RESULT 14
BG242597
LOCUS       BG242597                1116 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  602354010F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4482362 5',
            mRNA sequence.
ACCESSION   BG242597
VERSION     BG242597.1  GI:12752412
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1116)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10319  row: c  column: 03
            High quality sequence stop: 666.
FEATURES             Location/Qualifiers
     source          1..1116
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4482362"
                     /clone_lib="NCI_CGAP_Mam1"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="10 months, virgin"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
BASE COUNT      209 a    349 c    314 g    244 t
ORIGIN

  Query Match             15.9%;  Score 543.4;  DB 12;  Length 1116;
  Best Local Similarity   84.8%;  Pred. No. 1.2e-62;
  Matches  644;  Conservative    0;  Mismatches  111;  Indels    4;  Gaps    3;

Qy       1023 GCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCA 1082
              || ||||||||||||| |||||  ||||| ||||||||||||||| ||||||||||||||||
Db          9 GCCTGGCTTTCCGGAATCTGGGTACCCTGTTTCCCCGGCTGCAGCAGCTGTGCTGCCGCA 68

Qy       1083 TGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGA 1142
              |||| |||||||| ||||| ||||| |||||||||||||||||||||||||||||| ||||
Db         69 TGCCTCGCACCCTACGCCGACTCTTTGTGGCTGAGCTGTGCAGCTGGATGGCACTTATGA 128

Qy       1143 CCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAG 1202
              | ||||| ||||| |||||||| |||||||| |||||||||||||||| || |||||||
Db        129 CTTTCACACTGTTCTACACGGACTTCGTGGGAGAGGGGCTGTACCAGGGTGTACCCAGAG 188

Qy       1203 CTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGG 1262
              | ||||| |||||||||||||||||||||||||||||||||| |||| ||||||||||||
Db        189 CCGAGCCAGGCACCGAGGCCCGGAGACACTATGATGAAGGCATTCGAATGGGCAGCCTGG 248

Qy       1263 GGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGC 1322
              |||| |||||||||||| |||||||||||||||||||| ||||||||||||| ||||||| |
Db        249 GGCTCTTCCTGCAGTGTGCCATCTCCCTGGTCTTCTCCCTGGTCATGGACAGGCTGGTAC 308

Qy       1323 AGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCG 1382
              ||   ||||||||| ||  |||||||| ||||||||||       | || |||||||||||||
Db        309 AGAAGTTCGGCACACGGTCAGTCTATCTGGCCAGTGTGATGACCTTTCCTGTGGCTGCCG 368

Qy       1383 GTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGT 1442
              |||||| ||||||||||||||| ||||  || |||||||| ||||| |||||||||||||
Db        369 CTGCCACCTGCCTGTCCCACAGCGTGGTGGTAGTGACAGCCTCAGCTGCCCTCACCGGGT 428

Qy       1443 TCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGA 1502
              ||||||||| |||  |||||||||||||| ||||| || |||||||||||||||||| ||||
Db        429 TCACCTTCTCGGCCTTGCAGATCCTGCCTTACACGCTCGCCTCCCTCTACCACCGTGAGA 488

Qy       1503 AGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCC 1562
              ||||||||||||||||||||||||||||||| |||||||| ||| ||  |||||||||||
Db        489 AGCAGGTGTTCCTGCCCAAATACCGAGGGGACGCTGGAGGTAGCAGCGGTGAGGACAGCC 548

Qy       1563 TGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGG 1622
              ||  ||||||||| ||||||||||||||||||||| |||||| ||| || |||||||||||||
Db        549 AGACAACCAGCTTCTTGCCAGGCCCTAAGCCAGGAGCTCTCTTTCCCAATGGACACGTGG 608

Qy       1623 GTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTG 1682
              | ||||   |    |  | | ||    ||  |  | || ||||| ||||||||  ||| |
Db        609 GCTCTGGCAACCAGCGGATCCT--GGCCCTCCACTGCACTCTGTGGGGCCTCTTGCTGCG 666

Qy       1683 ATGTCTCCGTACGTGT-GGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGG 1741
              |||| ||| | || || ||||||||||||||  |    || |  ||| || ||||||||
Db        667 ATGTTTCCATGCGAGTGGGTGGTGGGTGAGCCACCTTGGGGCCCGGTTGTCCCGGGCCGG 726

Qy       1742 GGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTC 1780
              ||||| ||||||| ||| ||||| ||||    ||||||||
Db        727 GGCATTTGCCTGG-CCTGGCCATTCTGGCCGGTGCCTTC 764


RESULT 15
BG828841
LOCUS       BG828841                1035 bp    mRNA    linear   EST 22-MAY-2001
DEFINITION  602752813F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4905674 5',
            mRNA sequence.
ACCESSION   BG828841
VERSION     BG828841.1  GI:14176428
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1035)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1807  row: i  column: 03
```

```
                     /clone="IMAGE:5031771"
                     /clone_lib="NCI_CGAP_Mam3"
                     /tissue_type="tumor, gross tissue"
                     /dev_stage="10 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
                     Reference for transgenic model: Xu et al., Nature Genetics
                     22, 37-43 (1999)."
BASE COUNT      133 a    265 c    245 g    215 t
ORIGIN

  Query Match             17.2%;  Score 588;  DB 13;  Length 858;
  Best Local Similarity   83.0%;  Pred. No. 1.8e-68;
  Matches  694;  Conservative    0;  Mismatches  140;  Indels    2;  Gaps    2;

Qy        648 TGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTC 707
Db         14 TTGCCCTGCTGATCTTGGGAGTGGGGCTGCTGGACTTTTGTGGCCAGGTGTGCTTTACTC 73

Qy        708 CACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACT 767
Db         74 CATTGGAGGCCTTACTCTCCGACCTCTTCCGGGACCCAGACCACTGCCGCCAAGCCTTCT 133

Qy        768 CTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTG 827
Db        134 CTGTCTACGCCTTCATGATCAGCCTTGGGGGCTGCCTGGGCTACCTCTTACCTGCCATTG 193

Qy        828 ACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCC 887
Db        194 ACTGGGACACCAGCGTTCTGGCCCCCTACCTGGGTACTCAGGAAGAATGCCTCTTTGGCC 253

Qy        888 TGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAG 947
Db        254 TCCTCACCCTCATTTTCCTCATCTGCATGGCAGCCACTCTGTTTGTGACGGAGGAGGCAG 313

Qy        948 CGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCT 1007
Db        314 TACTGGGCCCACCCGAGCCGGCAGAAGGGTTGTTGGTCTCTGCCGTGTCGCGCCCGATGCT 373

Qy       1008 GTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACC 1067
Db        374 GCCCATGCCACGTTGGCCTGGCTTTCCGGAATCTGGGTACCCTGTTTCCCCGGCTGCAGC 433

Qy       1068 AGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCT 1127
Db        434 AGCTGTGCTGCCGCATGCCTCGCACCCTACGCCGACTCTTTGTGGCTGAGCTGTGCAGCT 493

Qy       1128 GGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACC 1187
Db        494 GGATGGCACTTATGACTTTCACACTGTTCTACACGGACTTCGTGGGAGAGGGGCTGTACC 553

Qy       1188 AGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTC 1247
Db        554 AGGGTGTACCCAGAGCCGAGCCAGGCACCGAGGCCCGGAGACACTATGATGAAGGCATTC 613

Qy       1248 GGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCA 1307
Db        614 GAATGGGCAGCCTGGGGCTCTTCCTGCAGTGTGCCATCTCCCTGGTCTTCTCCCTGGTCA 673

Qy       1308 -TGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCT 1366
Db        674 TTGGACAGGCTGGTACAGAAGTTCGGCACACGGTCAGTCTATCTGGCCAGTGTGATGACC 733

Qy       1367 TTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCA 1426
Db        734 TTTCCTGTGG-TGCCGCTGCACTGCCTGTCCCACAGCGTGGTGGTAAGTGACAAGCCTCA 792

Qy       1427 GCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGC 1482
Db        793 AGTGGCCTCACCGGGTTCACTTTCTCGGCCTTGAAGATCTGGCTTTAAAGGTCGGC 848
```

```
RESULT 13
BM914562
LOCUS       BM914562                1060 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6615475 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:5480245
            5', mRNA sequence.
ACCESSION   BM914562
VERSION     BM914562.1  GI:19364941
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1060)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Mark Watson
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2002  row: m  column: 14
            High quality sequence stop: 485.
FEATURES             Location/Qualifiers
     source          1..1060
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5480245"
                     /clone_lib="NIH_MGC_113"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: spleen; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed  by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of  California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
BASE COUNT      282 a    294 c    314 g    169 t      1 others
ORIGIN

  Query Match             16.5%;  Score 564;  DB 14;  Length 1060;
  Best Local Similarity   98.3%;  Pred. No. 2.3e-65;
  Matches  570;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1536 CTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTG 1595
Db          1 CTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTG 60

Qy       1596 GAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCAC 1655
Db         61 GAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCAC 120

Qy       1656 CCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCA 1715
Db        121 CCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCA 180

Qy       1716 CCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTG 1775
Db        181 CCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTG 240

Qy       1776 CCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCC 1835
Db        241 CCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCC 300

Qy       1836 AGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTG 1895
```

```
Db        481 CCACTCCCCTCTACTCTCTAGGACTGGGCTGATGAANGCACTGCCCCAAATTTTCCCT 540

Qy       2998 ACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCT-GTTTGGAGC 3056
              ||||||||||||||||||||||||||||||||||| || ||  |||||| |||||| |
Db        541 ACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACAAGTTCACCAACCCTGGTTTGGAAC 600

Qy       3057 TACTGC--AGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCC 3114
              ||||||   |  ||  |  | |||||||||| ||||||||||||||  | | |   ||||
Db        601 TACTGCCAGGACCCGAAGCCCCAAAGTGCGGGTTCCCAAGCCTTTGGGCCTTTTCACCCC 660

Qy       3115 C-----AGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGC 3169
              |     ||| ||| |  |    | ||| || | | |  |||| ||  | | |||||  |
Db        661 CCCAGAAGAATATCTGGGCTTTGGGAAACCCCCCCCGAAAATCTCGGGGACCCCCCCC 720

Qy       3170 CTGAGCTAA 3178
              |||  || |
Db        721 CTGCCCTGA 729
```

RESULT 11
BG469520
LOCUS   BG469520   894 bp   mRNA   linear   EST 21-MAR-2001
DEFINITION 602532833F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4660496 5',
            mRNA sequence.
ACCESSION   BG469520
VERSION   BG469520.1 GI:13401795
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 894)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: NIH Intramural Sequencing Center
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1458 row: a column: 09
            High quality sequence stop: 671.
FEATURES   Location/Qualifiers
     source   1..894
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4660496"
                     /clone_lib="NIH_MGC_15"
                     /tissue_type="adenocarcinoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G). Size-selected >500bp for average
                     insert size 1.8kb. Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)"
BASE COUNT   173 a   272 c   264 g   185 t
ORIGIN

  Query Match   17.4%; Score 593; DB 12; Length 894;
  Best Local Similarity 97.7%; Pred. No. 3.9e-69;
  Matches 644; Conservative   0; Mismatches   10; Indels   5; Gaps   4;

```
Qy       1450 CTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGT 1509
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGT 61

Qy       1510 GTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGAC 1569
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGAC 121

Qy       1570 CAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGG 1629
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGG 181

Qy       1630 AGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTC 1689
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 AGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTC 241

Qy       1690 CGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCT- 1748
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTC 301

Qy       1749 GCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGT 1808
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 GCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGT 361

Qy       1809 TTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCA- 1867
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAG 421

Qy       1868 GGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTG 1927
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 GGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTG 481

Qy       1928 GCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGG 1987
              |||||||||||||||||||||||||||||||||| ||||||||||||||| |||||||||
Db        482 GCCAAATACTCAGCGTAGAAAACTTCCAGCACATAGGGGTGGAGGGCCTGGCTCACTGGG 541

Qy       1988 TCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCC-GGGCTGGCCGCCA--GTTTCT 2044
              |||||||||||||||||||||||||||||||||||||| |||||||||||||   ||
Db        542 TCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGGCTGGCCGCCAAGTTTCTG 601

Qy       2045 GTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCT 2103
               | |||| ||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        602 TTGGCTGGCAAAGTAATGTGGCTCTCTGCTGACACCCTGTGCTGCTGAGGTGCGTAGCT 660
```

RESULT 12
BI107873
LOCUS   BI107873   858 bp   mRNA   linear   EST 26-JUN-2001
DEFINITION 602901816F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5031771 5',
            mRNA sequence.
ACCESSION   BI107873
VERSION   BI107873.1 GI:14558766
KEYWORDS   EST.
SOURCE   house mouse.
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 858)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11088 row: o column: 04
            High quality sequence stop: 810.
FEATURES   Location/Qualifiers
     source   1..858
                     /organism="Mus musculus"
                     /strain="129,C57BL/6J,FVB/N"
                     /db_xref="taxon:10090"

```
Qy       1468 GCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCG 1527
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCG 61

Qy       1528 AGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCC 1587
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 AGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCC 121

Qy       1588 TAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCC 1647
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 TAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCC 181

Qy       1648 ACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGG 1707
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 ACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGG 241

Qy       1708 TGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCT 1767
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        242 TGAGCCCACCGAGGCCA-GGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCT 300

Qy       1768 GGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCA 1827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCA 360

Qy       1828 GCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCAT 1887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCAT 420

Qy       1888 TTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAA 1947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAA 480

Qy       1948 AACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGT 2007
              ||||||||||||| |||||||||||||||||||||||||||||||||||| |||||||||
Db        481 AACTTCCAGCACA-TGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCTGCTCCTGT 539

Qy       2008 TAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCT 2067
              |||||||||||||||| |||||||||||      ||      ||||||||||||||||||
Db        540 TAGCCCCATGGGGCTG-CGGGCTGGCCGC--AGTTTCTGTGCTGCCAAAGTAATGTGGCT 596

Qy       2068 CTCTGCT--GCCACCCTGTGCTGCTGAGGTGCGTAGC-TGCACAGCT-GGGGGCTGGGGC 2123
              |||||||    |  ||  |||||||||||||||||||| ||||||||| |||||||||||
Db        597 CTCTGCTTGCCACCCTGGTGCTGCTGAGGTGCGTAGCTTGCACAGCTGGGGGGCTGGGGC 656

Qy       2124 GTCCCT-CTCCTCTCTCCCCAGTCTCTAGGGCT-GCCTGACTGGAGGCCTTCCAAGGGGG 2181
              |||||| ||||||||||||||||| ||||||||| ||||||||||||||||||| |||||
Db        657 GTCCCTCCTCCTCTCTCCCCAGTTTCTAGGGCTCGCCTGACTGGAGGCCTTCCAGGGGGT 716

Qy       2182 TTTCAGTCTGGAC-TTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGAC 2240
               |          |||| |||||||||| |||||||| || ||  ||  |   | || |
Db        717 CTCAGTCTGGGACTTTATACAGGCAGGCCAGAAAGGTCCCTGGCCTGGATGGGGGGCCTC 776

Qy       2241 TCTGCAGGT 2249
               |  |  |||
Db        777 TGCGGTGGT 785
```

RESULT 10
BQ942028

```
LOCUS       BQ942028               875 bp    mRNA    linear   EST 21-AUG-2002
DEFINITION  AGENCOURT_8817111 Lupski_sciatic_nerve Homo sapiens cDNA clone
            IMAGE:6202754 5', mRNA sequence.
ACCESSION   BQ942028
VERSION     BQ942028.1  GI:22357506
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 875)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13621  row: f  column: 03
            High quality sequence stop: 417.
FEATURES             Location/Qualifiers
     source          1..875
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6202754"
                     /clone_lib="Lupski_sciatic_nerve"
                     /sex="male"
                     /tissue_type="sciatic nerve"
                     /dev_stage="adult, 70 yr"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.87 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated  by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
BASE COUNT      196 a    260 c    201 g    217 t      1 others
ORIGIN

  Query Match             17.4%;  Score 593.2;  DB 14;  Length 875;
  Best Local Similarity   92.0%;  Pred. No. 3.7e-69;
  Matches  671;  Conservative    0;  Mismatches   49;  Indels    9;  Gaps    4;

Qy       2459 GAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTG 2518
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTG 60

Qy       2519 CTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCAT 2578
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCAT 120

Qy       2579 CACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCC 2638
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCC 180

Qy       2639 ATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAAT 2698
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAAT 240

Qy       2699 CAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGG 2758
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGG 300

Qy       2759 GTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATA 2818
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATA 360

Qy       2819 ATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAG 2878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAG 420

Qy       2879 GTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTT-CCCCCCACTT 2937
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||| ||
Db        421 GTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCCATTT 480

Qy       2938 CCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCT 2997
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||| ||||
```

```
            IMAGE:6204253 5', mRNA sequence.
ACCESSION   BQ950805
VERSION     BQ950805.1  GI:22366283
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 959)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13625  row: d  column: 14
            High quality sequence start: 2
            High quality sequence stop: 490.
FEATURES             Location/Qualifiers
     source          1..959
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6204253"
                     /clone_lib="Lupski_sciatic_nerve"
                     /sex="male"
                     /tissue_type="sciatic nerve"
                     /dev_stage="adult, 70 yr"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.87 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated  by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
BASE COUNT      139 a    312 c    308 g    200 t
ORIGIN

  Query Match             17.7%;  Score 604.2;  DB 14;  Length 959;
  Best Local Similarity   97.9%;  Pred. No. 1.3e-70;
  Matches  612;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       1614 GACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCT 1673
              ||   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GAACCGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCT 62

Qy       1674 CTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTC 1733
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 CTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTC 122

Qy       1734 CGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGG 1793
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 CGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGG 182

Qy       1794 TGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATA 1853
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 TGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATA 242

Qy       1854 TGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTG 1913
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 TGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTG 302

Qy       1914 ACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGG 1973
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGG 362

Qy       1974 CCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGC 2033
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 CCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGC 422

Qy       2034 CGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAG 2093
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 CGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAG 482

Qy       2094 GTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGG 2153
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 GTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGG 542

Qy       2154 CTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAA 2213
              |||||||||||||||||||||||||||| ||||||||||||||||||   |  |||  ||
Db        543 CTGCCTGACTGGAGGCCTTCCAAGGGGGCTTCAGTCTGGACTTATACCACGGAGGCCGAA 602

Qy       2214 GGGCTCCATGCACTGGAATGCGGGG 2238
              |||||||||| | ||||||| ||||
Db        603 GGGCTCCATGGATTGGAATGGGGGG 627


RESULT 9
BF972601
LOCUS       BF972601                 850 bp    mRNA    linear   EST 22-JAN-2001
DEFINITION  602243025F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4331407 5',
            mRNA sequence.
ACCESSION   BF972601
VERSION     BF972601.1  GI:12339816
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 850)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1196  row: a  column: 08
            High quality sequence stop: 675.
FEATURES             Location/Qualifiers
     source          1..850
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4331407"
                     /clone_lib="NIH_MGC_46"
                     /tissue_type="leiomyosarcoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Size-selected >500bp for average insert size
                     1.8kb. Library constructed by Ling Hong in the laboratory
                     of Gerald M. Rubin (University of California, Berkeley)
                     using  ZAP-cDNA synthesis kit (Stratagene) and Superscript
                     II RT (Life Technologies). Note: this is a NIH_MGC
                     Library."
BASE COUNT      128 a    266 c    261 g    195 t
ORIGIN

  Query Match             17.4%;  Score 594.6;  DB 12;  Length 850;
  Best Local Similarity   92.9%;  Pred. No. 2.4e-69;
  Matches  733;  Conservative    0;  Mismatches   44;  Indels   12;  Gaps   10;
```

```
Db         62 CGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGC 121

Qy       1678 CTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGG 1737
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        122 CTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCA-GGTGGTTCCGGG 180

Qy       1738 CCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGC 1797
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGC 240

Qy       1798 CCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGT 1857
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGT 300

Qy       1858 GTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAA 1917
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAA 360

Qy       1918 GAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTG 1977
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTG 420

Qy       1978 C-CTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGC 2036
              | ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        421 CACTCACTGGGTCCCAGCTCCCCGATCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGC 480

Qy       2037 CAGTTTCTGTTGCTGCCAAAGTAATGTGG-CTCTCTGCTGCCACCCTGTGCTGCTGAGGT 2095
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        481 CAGTTTCTGTTGCTGCCAAAGTAATGTGGCCTCTCTGCTGCCACCCTGTGCTGCTGAGGT 540

Qy       2096 GCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCT 2155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCT 600

Qy       2156 GCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGG 2215
              ||||||||||||||||||||||||||| |||||||||||||||||| |||||||||||||
Db        601 GCCTGACTGGAGGCCTTCCAAGGGGGTATCAGTCTGGACTTATACA-GGAGGCCAGAAGG 659

Qy       2216 GCTCCATGCACTGGAATGCGGGGACTCTGC-AGGTGGATTACCCAGGCTCAGGGTTAACA 2274
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db        660 GCTCCATGCACTGGAATGCGGGGACTCTGCAAGGTGGATTACCCAGGCTCAGGGTTAACA 719

Qy       2275 GCTA--GCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAG 2319
              ||||   |  |   ||||||||||||||||   | || |   |||||
Db        720 GCTAAGCCTCCTAGGTTGAGACACACCTAGGAGACGGGTCACGGGAG 766
```

```
RESULT 7
BE867241
LOCUS       BE867241                 718 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601442309F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3846411 5',
            mRNA sequence.
ACCESSION   BE867241
VERSION     BE867241.1  GI:10316017
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 718)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9559  row: a  column: 04
            High quality sequence stop: 693.
FEATURES             Location/Qualifiers
     source          1. .718
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3846411"
                     /clone_lib="NIH_MGC_65"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.  Library constructed by Life
                     Technologies. "
BASE COUNT      112 a    227 c    210 g    168 t      1 others
ORIGIN

  Query Match            19.4%;  Score 660.2;  DB 12;  Length 718;
  Best Local Similarity   97.7%;  Pred. No. 6.2e-78;
  Matches  690;  Conservative    0;  Mismatches   14;  Indels    2;  Gaps    2;

Qy       1300 TCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGT 1359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGT 60

Qy       1360 GGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGAC 1419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGAC 120

Qy       1420 AGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACT 1479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACT 180

Qy       1480 GGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGG 1539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGG 240

Qy       1540 AGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGC 1599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGC 300

Qy       1600 TCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGC 1659
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGC 360

Qy       1660 GCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGA 1719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGA 420

Qy       1720 GGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTT 1779
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTT 480

Qy       1780 CCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTC 1839
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTC 540

Qy       1840 -TGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCA-TTTACTTTGCT 1897
               |||||||  ||||||||||||| ||||||||||||||||||||||||| |||||||||||
Db        541 TTGTCACTNGCTATATGGTGTCTTGCGCAGGCCTGGGTCTGGTCGCCATTTTACTTTGCT 600

Qy       1898 ACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGC 1957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGC 660

Qy       1958 ACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTC 2003
              ||||||||||||||||||||      ||||||||  |  ||||  ||
Db        661 ACATTGGGGTGGAGGGCCTGCTCACTGGGGTCCCAAGTTCCCGTTC 706
```

```
RESULT 8
BQ950805
LOCUS       BQ950805                 959 bp    mRNA    linear   EST 21-AUG-2002
DEFINITION  AGENCOURT_8842232 Lupski_sciatic_nerve Homo sapiens cDNA clone
```

```
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: Dr. James R. Lupski
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM13628  row: b  column: 19
           High quality sequence stop: 579.
FEATURES             Location/Qualifiers
     source          1. .916
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6205362"
                     /clone_lib="Lupski_sciatic_nerve"
                     /sex="male"
                     /tissue_type="sciatic nerve"
                     /dev_stage="adult, 70 yr"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.87 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated  by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
BASE COUNT       215 a    264 c    193 g    244 t
ORIGIN

  Query Match             20.6%;  Score 703.2;  DB 14;  Length 916;
  Best Local Similarity   98.8%;  Pred. No. 1.2e-83;
  Matches  751;  Conservative    0;  Mismatches    3;  Indels    6;  Gaps    4;

Qy       2277 TAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCA 2336
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCA 60

Qy       2337 CCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATG 2396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATG 120

Qy       2397 GGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGG 2454
              |||||||||||||||||||||||||||||||||||||||||||||  | |||||||||||
Db        121 GGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGAAAGTTATTTGTAGG 180

Qy       2455 GGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTT 2514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTT 240

Qy       2515 TTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTG 2574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTG 300

Qy       2575 CCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGA 2634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGA 360

Qy       2635 ATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAA 2694
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAA 420

Qy       2695 CAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCC 2754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCC 480

Qy       2755 TGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAAT 2814
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAAT 540

Qy       2815 GATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCT 2874
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCT 600

Qy       2875 TCAGGTCTCAACGGC-TTCCCTAACCACCCCTCTTCTCTTGGCCCAG-CCTGGTTCCCCC 2932
              ||||||||||||||| ||||||||||||||||||||||||||||||| ||||||||||||
Db        601 TCAGGTCTCAACGGCTTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCCTGGTTCCCCC 660

Qy       2933 CACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTT 2992
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTT 720

Qy       2993 CCCCTACCCCCAAC--TTTCCCCTACCCCCAACTTTCCCC 3030
              ||||||||||||||  |||||| |||||| ||||||||||
Db        721 CCCCTACCCCCAACTTTTTCCCTTACCCCAAACTTTCCCC 760


RESULT 6
BG469487
LOCUS       BG469487                 800 bp    mRNA    linear   EST 21-MAR-2001
DEFINITION  602532993F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4660703 5',
            mRNA sequence.
ACCESSION   BG469487
VERSION     BG469487.1  GI:13401762
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 800)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: NIH Intramural Sequencing Center
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1458  row: i  column: 24
            High quality sequence stop: 741.
FEATURES             Location/Qualifiers
     source          1. .800
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4660703"
                     /clone_lib="NIH_MGC_15"
                     /tissue_type="adenocarcinoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)"
BASE COUNT       135 a    245 c    239 g    181 t
ORIGIN

  Query Match             19.6%;  Score 667.4;  DB 12;  Length 800;
  Best Local Similarity   97.0%;  Pred. No. 6.6e-79;
  Matches  744;  Conservative    0;  Mismatches   16;  Indels    7;  Gaps    6;

Qy       1558 CAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACA 1617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 CAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACA 61

Qy       1618 CGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGC 1677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy       1022 CGCTTGGCTTTCCGGAA--CCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCC 1079
              |||||||||||||||||  |  |||||||||||||||||||||||||||||||||||||
Db        660 CGCTTGGCTTTCCGGAAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCC 719

Qy       1080 GCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGG--ATGGCACT 1137
               |     ||||  |  ||||||||||||||||||||||||||||||||||   |  |||
Db        720 GGCATGCCCGGAACCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGCATGGCAACT 779

Qy       1138 CATGACCTTCACGCTGTTTTACACGG--ATTTCGTGGCGAGGGGCTGT-ACCAGGGCGT 1194
              |||||||||||  ||||| |||  ||    |    ||||||| ||||||| ||| |||||
Db        780 CATGACCTTCAGCCTGTTCTACCCGGCATCTCGCTGGGCGACGGGCTGTCCCCACGGCGT 839

Qy       1195 GCCCAGAGCTGAGCCGGGCACC-GAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGG 1253
                 |||||||||||||||||||| || |||||||||  ||||| |||| |||| || || |
Db        840 TGCCAGAGCTGAGCCGGGCACCGGAAGCCCGGAGAACCTATGCTGAACGCGTCCGCAT-G 898

Qy       1254 GCAGCCTGGGGCTGTTCCTGCAGTGC 1279
              || ||||||| |||| |    ||| ||
Db        899 GCCGCCTGGGCCTGTCCTGACAGCGC 924
```

RESULT 4
BM915082
LOCUS   BM915082   1025 bp   mRNA   linear   EST 12-MAR-2002
DEFINITION AGENCOURT_6702317 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5481218
            5', mRNA sequence.
ACCESSION   BM915082
VERSION   BM915082.1 GI:19365461
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1025)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2005 row: f column: 03
            High quality sequence start: 35
            High quality sequence stop: 657.
FEATURES   Location/Qualifiers
     source   1..1025
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5481218"
                     /clone_lib="NIH_MGC_41"
                     /tissue_type="amelanotic melanoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
BASE COUNT   237 a   273 c   252 g   263 t
ORIGIN

  Query Match   22.0%; Score 748.6; DB 14; Length 1025;
  Best Local Similarity   96.0%; Pred. No. 1.2e-89;
  Matches 802; Conservative   0; Mismatches   24; Indels   9; Gaps   3;

```
Qy       1799 CCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTG 1858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 CCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTG 71

Qy       1859 TCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAG 1918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 TCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAG 131

Qy       1919 AGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGC 1978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 AGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGC 191

Qy       1979 CTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCA 2038
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 CTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCA 251

Qy       2039 GTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCG 2098
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 GTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCG 311

Qy       2099 TAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCC 2158
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 TAACTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCC 371

Qy       2159 TGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCT 2218
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 TGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCT 431

Qy       2219 CCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTA 2278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 CCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTA 491

Qy       2279 GCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACC 2338
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 GCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACC 551

Qy       2339 TGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGG 2398
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        552 TGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGG 611

Qy       2399 AGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGG 2456
              |||||||||||||||||||||||||||||||||||||||||||    | ||||||||||||
Db        612 AGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGAAAGTTATTTGTAGGGG 671

Qy       2457 AAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTT 2516
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        672 AAGAGTCCTGAGGGGCAACACCCAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTT 731

Qy       2517 TGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC---TTCTGTT 2573
              ||||||||||||||||||||||||||||||||||||||||||      | ||    |   |||
Db        732 TGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCCTGCTGGCCCCTTCTGGTT 791

Qy       2574 GCCATCACAGAGACACAGGCATTT----AAATATTTAACTTATTTATTTAACAAA 2624
              ||| || || | |      ||  | |    ||| ||||||||||||||||||||||
Db        792 GCCTTCCCACAAAACACGGGGTATTTAAAAAAATTTAACTTATTTATTTAACAAA 846
```

RESULT 5
BQ950912
LOCUS   BQ950912   916 bp   mRNA   linear   EST 21-AUG-2002
DEFINITION AGENCOURT_8754471 Lupski_sciatic_nerve Homo sapiens cDNA clone
            IMAGE:6205362 5', mRNA sequence.
ACCESSION   BQ950912
VERSION   BQ950912.1 GI:22366390
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 916)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.

```
Db        121 GACTTGGCCAAATACTCATCGTACAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTC 180

Qy       1982 ACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTT 2041
Db        181 ACTGGGTCCCAGCTCCCCGCTCCTGATAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTT 240

Qy       2042 TCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAG 2101
Db        241 TCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGACGTGCGTAG 300

Qy       2102 CTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGA 2161
Db        301 CTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGA 360

Qy       2162 CTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCA 2221
Db        361 CTGGAGGCCTTCCAAGGGGGATTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCA 420

Qy       2222 TGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCC 2281
Db        421 TGCACTGGAATGCGGGGACTCTGCATGTGGATTACCCATGCTCAAGGTTAACAGCTAGCC 480

Qy       2282 TCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGG 2341
Db        481 TCCTAGTTGAGACACACCTAGAGAATGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGG 540

Qy       2342 TTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGT 2401
Db        541 TTTCCCATCTCTAAACCCCTTAACCTGCAACTTCGTTTAATGTAGCTCTTGCATGGGAGT 600

Qy       2402 TTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAG 2459
Db        601 TTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGAAAGTTATTTGTATGGGAAG 660

Qy       2460 AGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGC 2519
Db        661 AGTCCTGAGGGGCAACACACTAGAACCAGGTCCCCTCAACCCACAGCACTGTCTTTATGC 720

Qy       2520 TGATCCAC-CCCCCTCTTACCTTTTATCAGGATGTGG-CCTGTTGGTCCTTCTGTTGCCA 2577
Db        721 TGATCCACGCCCCCTCTTTACTTTTATCAAGATGTGGCCCTGTTGGTCCTTCTGTTGCCA 780

Qy       2578 TCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAA-CAAAGTAGAAGGGAAT 2636
Db        781 TCACAGAGACACAGGCATTTAATATTTAACTTATTTATTTAANCAAAGTATAAAGGAAT 840

Qy       2637 CCATTGCTAGCTTTT--CTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 2686
Db        841 CCCTTGCTACCTTTTTCCGAGATGGAGGACAAAAATTTGGGTAAGGGAAGGG 892
```

```
RESULT 3
BG469889
LOCUS       BG469889          946 bp    mRNA    linear   EST 21-MAR-2001
DEFINITION  602533345F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4661075 5',
            mRNA sequence.
ACCESSION   BG469889
VERSION     BG469889.1  GI:13402164
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 946)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: NIH Intramural Sequencing Center
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1459  row: i  column: 12
            High quality sequence stop: 798.
FEATURES             Location/Qualifiers
     source          1..946
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4661075"
                     /clone_lib="NIH_MGC_15"
                     /tissue_type="adenocarcinoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)"
BASE COUNT      120 a    331 c    296 g    199 t
ORIGIN

  Query Match             22.0%;  Score 749.6;  DB 12;  Length 946;
  Best Local Similarity   93.1%;  Pred. No. 9.2e-90;
  Matches  862;  Conservative    0;  Mismatches   54;  Indels   10;  Gaps    7;

Qy        362 GGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTG 421
Db          1 GGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTG 60

Qy        422 GGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTC 481
Db         61 GGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTC 120

Qy        482 TGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGG 541
Db        121 TGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGG 180

Qy        542 CCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCC 601
Db        181 CCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCC 240

Qy        602 GGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATC 661
Db        241 GGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATC 300

Qy        662 CTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTG 721
Db        301 CTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTG 360

Qy        722 CTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTC 781
Db        361 CTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTC 420

Qy        782 ATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGT 841
Db        421 ATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGT 480

Qy        842 GCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATC 901
Db        481 GCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATC 540

Qy        902 TTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACC 961
Db        541 TTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGCAGGAGGCAGCGCTGGGCCCACC 600

Qy        962 GAGCCAGCAGAAGGGCTGTCGGCCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCC 1021
Db        601 GAGCCAGCAGAAGGGCTGTCGG-CCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCC 659
```

```
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6198823"
                     /clone_lib="Lupski_sciatic_nerve"
                     /sex="male"
                     /tissue_type="sciatic nerve"
                     /dev_stage="adult, 70 yr"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.87 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated  by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
BASE COUNT      129 a    286 c    294 g    176 t
ORIGIN

  Query Match             23.0%;  Score 784.2;  DB 14;  Length 885;
  Best Local Similarity   95.1%;  Pred. No. 2.7e-94;
  Matches  842;  Conservative    0;  Mismatches   38;  Indels    5;  Gaps    3;

Qy         26 CCGGGTGACAGCCGCGCGCCTCGGCCAGGATCTGAGTGATGAGACGTGTCCCCACTGAGG 85
              ||  | | | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCACGCGTCCGCCGCGCGCCTCGGCCAGGATCTGAGTGATGAGACGTGTCCCCACTGAGG 60

Qy         86 TGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAGCTGGACCGGCACCAAAGGGCTG 145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGCCCCACAGCAGCAGGTGTTGAGCATGGGCTGAGAAGCTGGACCGGCACCAAAGGGCTG 120

Qy        146 GCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGCGGCAGCAAGGAGGAGAGGCCGC 205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGCGGCAGCAAGGAGGAGAGGCCGC 180

Qy        206 AGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAGTGCCTGAACGGCCCCCTGAGCC 265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAGTGCCTGAACGGCCCCCTGAGCC 240

Qy        266 CTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGG 325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGG 300

Qy        326 AAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCA 385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCA 360

Qy        386 GGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACC 445
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACC 420

Qy        446 ATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCC 505
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCC 480

Qy        506 AGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTG 565
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTG 540

Qy        566 GGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGC 625
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGC 600

Qy        626 CCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTC 685
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTC 660

Qy        686 TGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGA-CCC 744
              ||||||| |||||||||||||||||||||||||||||||||||| ||||||||||| |||
Db        661 TGTGGCCCGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGAACTCTTCCGGGACCCC 720

Qy        745 GGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTC--TTGGGGGCTGC 802
              ||||||||||||||||||||||||||||||||||||||||||| |||  | ||||   ||
Db        721 GGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCCGTCCTTGGGGGCTGGC 780

Qy        803 CTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGC 862
              | ||||||||||||||   | ||| |||||  ||||| | ||||| |||||| ||||||
Db        781 CGGGGCTACCTCCTGCCCTGCCTTGGCTGGGGAACCAGGGGCCTGGGCCCCTAACTGGGC 840

Qy        863 A--CCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCC 905
              |  ||||| ||||| ||| ||||||    |  | ||| |    || |
Db        841 AACCCCAGAAGGAGGGCCCCTTTGGGGCGGCCCCCCTCATTTC 885


RESULT 2
BM915527
LOCUS       BM915527                1063 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6701857 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5481717
            5', mRNA sequence.
ACCESSION   BM915527
VERSION     BM915527.1  GI:19365906
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1063)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2006  row: j  column: 22
            High quality sequence start: 131
            High quality sequence stop: 308.
FEATURES             Location/Qualifiers
     source          1..1063
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5481717"
                     /clone_lib="NIH_MGC_41"
                     /tissue_type="amelanotic melanoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
BASE COUNT      246 a    283 c    261 g    272 t      1 others
ORIGIN

  Query Match             22.5%;  Score 767.2;  DB 14;  Length 1063;
  Best Local Similarity   95.0%;  Pred. No. 4.2e-92;
  Matches  847;  Conservative    0;  Mismatches   38;  Indels    7;  Gaps    5;

Qy       1802 TCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCT 1861
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCT 60

Qy       1862 GCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGC 1921
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGC 120

Qy       1922 GACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTC 1981
```

```
OM nucleic - nucleic search, using sw model

Run on:         June  9, 2003, 15:13:08 ; Search time 4204 Seconds
                                           (without alignments)
                                           13136.682 Million cell updates/sec

Title:          US-09-759-143-110
Perfect score:  3410
Sequence:       1 gggaaccagcctgcacgcgc..........aaaaaaataaaaaaaaaaaa 3410

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters:      32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       EST:*
                 1:  em_estba:*
                 2:  em_esthum:*
                 3:  em_estin:*
                 4:  em_estmu:*
                 5:  em_estov:*
                 6:  em_estpl:*
                 7:  em_estro:*
                 8:  em_htc:*
                 9:  gb_est1:*
                 10:  gb_est2:*
                 11:  gb_htc:*
                 12:  gb_est3:*
                 13:  gb_est4:*
                 14:  gb_est5:*
                 15:  em_estfun:*
                 16:  em_estom:*
                 17:  gb_gss:*
                 18:  em_gss_hum:*
                 19:  em_gss_inv:*
                 20:  em_gss_pln:*
                 21:  em_gss_vrt:*
                 22:  em_gss_fun:*
                 23:  em_gss_mam:*
                 24:  em_gss_mus:*
                 25:  em_gss_other:*
                 26:  em_gss_pro:*
                 27:  em_gss_rod:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	784.2	23.0	885	14	BQ948390	BQ948390 AGENCOURT
	2	767.2	22.5	1063	14	BM915527	BM915527 AGENCOURT
	3	749.6	22.0	946	12	BG469889	BG469889 602533345
	4	748.6	22.0	1025	14	BM915082	BM915082 AGENCOURT
	5	703.2	20.6	916	14	BQ950912	BQ950912 AGENCOURT
	6	667.4	19.6	800	12	BG469487	BG469487 602532993
	7	660.2	19.4	718	12	BE867241	BE867241 601442309
	8	604.2	17.7	959	14	BQ950805	BQ950805 AGENCOURT
	9	594.6	17.4	850	12	BF972601	BF972601 602243025
	10	593.2	17.4	875	14	BQ942028	BQ942028 AGENCOURT
	11	593	17.4	894	12	BG469520	BG469520 602532833
	12	588	17.2	858	13	BI107873	BI107873 602901816
	13	564	16.5	1060	14	BM914562	BM914562 AGENCOURT
	14	543.4	15.9	1116	12	BG242597	BG242597 602354010
	15	538.2	15.8	1035	12	BG828841	BG828841 602752813
	16	534.8	15.7	589	12	BG469586	BG469586 602533622
	17	525	15.4	715	12	BG122427	BG122427 602353324
	18	512.8	15.0	1667	14	BM912193	BM912193 AGENCOURT
	19	508.6	14.9	872	12	BG864609	BG864609 602798469
	20	499.6	14.7	786	12	BG174399	BG174399 602334219
	21	463	13.6	901	13	BI650119	BI650119 603296208
	22	462.6	13.6	482	9	AF109299	AF109299 AF109299
	23	459	13.5	969	12	BG173136	BG173136 602335411
c	24	456.8	13.4	491	9	AA631143	AA631143 nq75g08.s
	25	456.2	13.4	929	12	BF785813	BF785813 602112437
	26	455.2	13.3	592	10	AW412402	AW412402 uo78h02.y
	27	453	13.3	934	14	BQ934815	BQ934815 AGENCOURT
c	28	449.2	13.2	469	9	AI703348	AI703348 wd93b09.x
	29	438.6	12.9	549	12	BG081576	BG081576 H3066G04-
	30	435.4	12.8	448	10	BE674096	BE674096 7d74g04.x
	31	433.8	12.7	700	12	BF581244	BF581244 602100464
	32	419.6	12.3	700	10	BB610495	BB610495 BB610495
	33	419.2	12.3	759	12	BF789072	BF789072 602104930
	34	409.4	12.0	692	10	BB627667	BB627667 BB627667
	35	400.2	11.7	844	12	BG246497	BG246497 602360526
c	36	398.6	11.7	412	10	AW175665	AW175665 RC3-BT004
c	37	395.8	11.6	418	12	BF922235	BF922235 QV2-NT014
	38	391	11.5	430	9	AF109303	AF109303 AF109303
	39	389.4	11.4	630	13	BI145201	BI145201 602909395
	40	381.2	11.2	650	10	BB627844	BB627844 BB627844
	41	369.4	10.8	537	10	AW787124	AW787124 120845 MA
c	42	366	10.7	388	9	AA579486	AA579486 nf33g07.s
	43	359.2	10.5	906	12	BF680993	BF680993 602156279
c	44	353	10.4	370	10	AW135465	AW135465 UI-H-BI1-
	45	352.8	10.3	509	10	BB701488	BB701488 BB701488

ALIGNMENTS

```
RESULT 1
BQ948390
LOCUS       BQ948390                 885 bp    mRNA    linear   EST 21-AUG-2002
DEFINITION  AGENCOURT_8803099 Lupski_sciatic_nerve Homo sapiens cDNA clone
            IMAGE:6198823 5', mRNA sequence.
ACCESSION   BQ948390
VERSION     BQ948390.1  GI:22363868
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 885)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13611  row: b  column: 08
            High quality sequence stop: 615.
FEATURES             Location/Qualifiers
     source          1..885
```

```
Db     243 erProGlyGlyLysGlyGlu---------------------------MetGlyProAlaG 254

Qy    2025 GGCAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGGACCCAGTGAGGCAGGCCCTCCAC 1966
                |||                  |||      ||||||         ||||||
Db     254 lyIleProGlyAlaProGlyLeuMetGlyAlaArgGlyPro------ProGlyProAlaG 272

Qy    1965 CCCAATGTGCTGGAAGTTTTCTACGCTG----AGTATTTGGCCAAGTCGCTCTTGTCAAA 1910
                               ||||||      :::      |||   ::::::
Db     272 lyAlaAsnGlyAlaProGlyLeuArgGlyGlyAlaGlyGluProGlyLysAsn------- 289

Qy    1909 TACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCA---------------------G 1871
                          :::|||        |||   |||                     |
Db     290 --------------GlyAlaLysGlyGluProGlyProArgGlyGluArgGlyGluAlaG 305

Qy    1870 GCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCAT 1811
           ||:::         |||      :::   :::   |||      |||
Db     305 lyIleProGlyValProGlyAlaLysGlyGluAspGlyLysAspGlySerProGlyGluP 325

Qy    1810 AAACAGGGATGGGGCCACCTGGGACAGCAGGA---AGGCACTATCCAGGATGGCGA---- 1758
                      |||   ||||||   ||||||   |||      ||||||:::|||
Db     325 roGlyAlaAsnGlyLeuProGlyAlaAlaGlyGluArgGlyAlaProGlyPheArgGlyP 345

Qy    1757 -----GGTCCAGGCAGATGCCCCGGCCCGGAACCACCCTGGCCTCGGTGGGCTCACCCAC 1703
                ||||||         ||||||   :::   |||                  |||
Db     345 roAlaGlyProAsnGlyIleProGlyGluLysGlyProAlaGlyGluArgGlyAlaProG 365

Qy    1702 CACCACACGTACGGAGACATCACAGGCAGAGGCCCCGCAGAGC-------GCGGGTGGAG 1650
             |||         |||               |||   |||            ||||||
Db     365 lyProAlaGlyProArgGlyAlaAlaGlyGluProGlyArgAspGlyValProGlyGlyP 385

Qy    1649 GTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGGGAAGGGAGCTCCAGGCT 1590
             |||   |||   :::|||   :::|||      |||         |||   ||||||
Db     385 roGlyMetArgGlyMetProGlySerProGlyGlyProGlySerAspGlyLysProGly- 404

Qy    1589 TAGGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCTAGCACCTCCAGTGTCCC 1530
                ||||||             ||||||   |||
Db     405 --ProProGly-SerGlnGlyGluSerGlyArgPro------------------------ 415

Qy    1529 CTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGTAGAGGGAGGCCAGTGTGTAGG 1470
                          |||   |||         |||   ||||||:::|||
Db     416 ---------------GlyProProGlyProSerGlyProArgGlyGlnPro--------- 427

Qy    1469 GCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAGCTGTCACCACGG 1410
              |||      |||                                       |||
Db     428 ---GlyValMetGlyPhe------------------------------------ProGly 434

Qy    1409 CCACACTGTGGGACAGGCATGTGGCACCGGCAGCCACAGGGAAAGCTGCCACACTGGCCA 1350
           |||            |||                     ||||||
Db     435 ProLysGlyAsnAspGly---------------AlaProGlyLys--------------- 444

Qy    1349 AATAGACTGCTCGAGTGCCGAATCGCTGCACCAGCCGGTCCATGACCAGAGAGAAGACCA 1290
                                         |||   ||||||   |||:::      |||
Db     445 ------------AsnGlyGluArgGlyGlyProGlyGlyProGlyProGlnGlyProPro 460

Qy    1289 GGGAGATGGCGCACTGCAGGA---------------------ACAGCCCCAGGCTGCCCA 1251
           |||:::            |||                     |||   ||||||
Db     461 GlyLysAsnGlyGluThrGlyProGlnGlyProProGlyProThrGlyProGlyGlyAsp 480

Qy    1250 TCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGC 1191
                                   :::||||||      |||      |||
Db     481 Lys---------------GlyAspThrGlyProProGlyProGlnGlyLeuGlnGlyLeu 495

Qy    1190 CCTGGTACA--------------------------------------------------- 1182
           |||||||||
Db     496 ProGlyThrGlyGlyProProGlyGluAsnGlyLysProGlyGluProGlyProLysGly 515

Qy    1181 ---------------------------------------------GCCCCTCGCCCACGA 1167
                                                        ||||||      |||
Db     516 AspAlaGlyAlaProGlyAlaProGlyGlyLysGlyAspAlaGlyAlaProGlyGluArg 535

Qy    1166 AATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGCTGCACAGCTCAGCCACGA 1107
              |||                           |||            |||
Db     536 GlyPro---------------------------ProGlyLeuAlaGlyAlaProGlyLeu 546

Qy    1106 AGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACAGCTGGTGCAGCCGGGGAAGCAGGG 1047
           |||    |||||||||                       ||||||||||||      |||
Db     547 ArgGlyGlyAlaGlyProProGlyProGluGlyGlyLysGlyAlaAlaGlyProProGly 566

Qy    1046 CGCCCAGGTTCC------------------------------------------------ 1035
              ||||||:::
Db     567 ProProGlyAlaAlaGlyThrProGlyLeuGlnGlyMetProGlyGluArgGlyGlyLeu 586

Qy    1034 ---------------------------GGAAAGCCAAGCGGGCCCGGCATGGACAGCAGT 1002
                                      |||:::|||   |||||||||   |||
Db     587 GlySerProGlyProLysGlyAspLysGlyGluProGlyGlyProGlyAlaAspGlyVal 606

Qy    1001 GGGGCGACA------AGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCG 948
                             ||||||||||||                     ||||||
Db     607 ProGlyLysAspGlyProArgGlyProThr---------------------GlyProIle 619

Qy     947 CTGCCTCCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCA 888
              ||||||   |||                           |||:::   |||
Db     620 GlyProProGlyProAla------------------------GlyGlnProGlyAspLys 631

Qy     887 GGCCAAAGAGGCACTCCTCCTGGGTGCCCAGGT---AGGGGGCCAGGGCACTGGTGTCCC 831
           |||:::   |||      ||||||   ||||||      ||||||
Db     632 GlyGluGlyGlyAla---ProGlyLeuProGlyIleAlaGlyPro--------------- 645

Qy     830 AGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCATAGA 771
                             ||||||||||||                     |||
Db     646 ---------------ArgGlySerProGlyGlu------------------ArgGlyGlu 654

Qy     770 CAGAGTAGGCCTGGCGACAGTGGT---CCGGGT---CCCGGAAGAGGTCAGAGAGCAGGG 717
                    ||||||   :::|||   ||||||   ||||||:::            |||
Db     655 ThrGlyProProGlyProAlaGlyPheProGlyAlaProGlyGlnAsnGlyGluProGly 674

Qy     716 CCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGCCCAGGATGA 657
                    |||         ||||||   :::            |||   ||||||
Db     675 GlyLysGlyGluArgGlyAlaProGlyGluLysGlyGluGlyGlyPro---ProGlyVal 693

Qy     656 GCAGTGCCAGCTCCAGGG---------GCCTGGGATCCGGGCACAGCAGCCCTGCTAGCC 606
           |||   |||      |||         |||      ||||||         :::
Db     694 AlaGlyProProGlyGlySerGlyProAlaGlyProProGlyProGlnGlyValLysGly 713

Qy     605 AGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCAAGGACAGTGCCCAGATGA 546
                             ||||||||||||   |||   |||
Db     714 ------------------GluArgGlySerProGlyGlyPro------------------ 721

Qy     545 AGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTAGGAGCGGGA 486
                                      |||:::|||               ||||||
Db     722 ---------------------------GlyAlaAlaGlyPhe------ProGlyAlaArg 730

Qy     485 CACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCATGGTCATGAACTTCTCCTCTA 426
                 |||||||||                     |||   |||      :::|||
Db     731 GlyLeuProGlyPro---------------------ProGlySerAsnGlyAsnProGly 743

Qy     425 CCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCCAAACACACCTCCA 366
           ||||||   |||:::                                          |||
Db     744 ProProGlyProSerGlySer------------------------------------Pro 751

Qy     365 GGCCAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGTGCCGCAGCAGGCGGC 306
           |||:::         |||   |||      ||||||   :::||||||         |||
Db     752 GlyLys------AspGlyProProGlyProAlaGlyAsnThrGlyAlaPro------Gly 767

Qy     305 TCACCCACAGCCTCTGGACCATAGTGGG------------CCAGGCGGGTAGGGCTCAGG 258
           ||||||      |||||||||   |||              ||||||      ||||||
Db     768 SerProGlyValSerGlyPro-LysGlyAspAlaGlyGlnProGlyGluLysGlySerPr 787

Qy     257 GGGCCGTTCAGGCACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCA-------------- 212
            |||      |||  |||                       ||||||
Db     787 oGlyAlaGlnGlyProPro------------------GlyAlaProGlyProLeuGlyIl 801

Qy     211 -------------------GAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAACTGCCTAG 171
                                        |||   ||||||               |||||
Db     801 eAlaGlyIleThrGlyAlaArgGlyLeuAlaGlyProProGlyMetProGly---ProAr 820
```

```
Qy       170 GAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGCTTCTCAG----- 116
             |    ||||||        :::                |||  :::
Db       820 gGlySerProGlyProGlnGlyValLysGlyGluSerGlyLysProGlyAlaAsnGlyLe 840

Qy       115 ------------------------CCCATGCTCAACACCTGCTGCTGTGGGGCACCTCA 81
                                     |||        |||        |||        :
Db       840 uSerGlyGluArgGlyProProGlyProGln-GlyLeuProGlyLeuAlaGly---ThrA 859

Qy        80 GTGGGGACACGTCTCATCACTCAGATCCTGGC--------------------------- 49
             ::|||:::                :::||||||
Db       859 laGlyGluProGlyArgAspGlyAsnProGlySerAspGlyLeuProGlyArgAspGlyS 879

Qy        48 -------------------CGAGGCGCGCGGCTGTCACCCGGAGCC 22
                                |||||||        ||||||||||||
Db       879 erProGlyGlyLysGlyAspArgGlyGluAsnGlySerProGlyAla 894
```

RESULT 43
US-09-219-849-49
; Sequence 49, Application US/09219849
; Patent No. 6150081
; GENERAL INFORMATION:
; APPLICANT: VAN HEERDE, GEORGE V.
; APPLICANT: VAN RIJN, ALEXIS C.
; APPLICANT: BOUWSTRA, JAN B.
; APPLICANT: DE WOLF, FREDERIK A.
; APPLICANT: MOOBROEK, ANDREAS
; APPLICANT: WERTEN, MARC W.T.
; APPLICANT: WIND, RICHELE D.
; APPLICANT: VAN DEN BOSCH, TANJA J.
; TITLE OF INVENTION: SILVER HALIDE EMULSIONS WITH RECOMBINANT COLLAGEN
; TITLE OF INVENTION: SUITABLE FOR PHOTOGRAPHIC APPLICATION AND ALSO THE
; TITLE OF INVENTION: PREPARATION THEREOF
; FILE REFERENCE: 2728-2
; CURRENT APPLICATION NUMBER: US/09/219,849
; CURRENT FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 822
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: amino acid sequence
US-09-219-849-49

Alignment Scores:

Pred. No.:	6.84e-19	Length:	822
Score:	351.00	Matches:	244
Percent Similarity:	32.94%	Conservative:	37
Best Local Similarity:	28.60%	Mismatches:	328
Query Match:	5.64%	Indels:	245
DB:	4	Gaps:	43

US-09-759-143-110 (1-3410) x US-09-219-849-49 (1-822)

```
Qy      2131 AGAGGGACGCCCCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGC 2072
             ||||||    |||    ||||||                            ||||||
Db         8 ArgGlyLeuProGlyProProGlyAlaPro---GlyProGlnGlyPheGlnGlyProPro 26

Qy      2071 AGAGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGG 2012
               ||||||        |||    :::             |||         |||
Db        27 GlyGluProGlyGluProGlyGlySerGlyProMetGlyProArgGlyProProGlyPro 46

Qy      2011 GCTAACAGGAGCGGGGAGCTGGGA--------------------CCCAGTGAGGCAGGC 1973
                   ::::::|||:::    |||                    |||   |||   |||
Db        47 ProGlyLysAsnGlyAspAspGlyGluAlaGlyLysProGlyArgProGlyGluArgGly 66

Qy      1972 CCTCCA----------------------CCCCAATGTGCTGGAAGTTTTCTACGCTGA 1937
             ||||||                      |||       ||||||
Db        67 ProProGlyProGlnGlyAlaArgGlyLeuProGlyThrAlaGly--------------- 81

Qy      1936 GTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAG 1877
               :::   |||                    :::                ||||||
Db        82 LeuProGlyMetLysGlyHisArgGlyPheSerGlyLeuAspGlyAlaLysGlyAspAla 101

Qy      1876 ACCCAGGCCTGCGGCAGACACCATATAGGCAGTGA----CAGACTGGCTGAGCTGGACAA 1821
               |||    :::       ||||||    :::   |||    ||||||
Db       102 GlyProAlaGlyProLysGlyGluProGlySerProGlyGluAsnGlyAlaProGlyGln 121

Qy      1820 TGGAGCCCATAAACAGGGATG--------------------GGGCCACCTGGGACAGCA 1782
                 |||     |||:::                 ||||||||||||||||||
Db       122 MetGlyPro---ArgGlyLeuProGlyGluArgGlyArgProGlyProProGlyThrAla 140

Qy      1781 GGAAGGCACTATCCAGGATGGCGAGGTCCAGGCAGATGCCCCGGCCCGGAACCACCCTGG 1722
             |||          ||||||  |||  |||         |||  ||||||  |||
Db       141 GlyAlaArgGly-AsnAspGlyAlaValGlyAlaAlaGlyPro-ProGlyProThr---G 159

Qy      1721 CCTCGGTGGGCTCACCCACCACCACACGTACGGAGACATCACAGGCAGAGGCC---CCGC 1665
                  |||   |||                     ::::::   ||||||   ||||
Db       159 lyProThrGlyProProGlyPheProGlyAlaValGlyAlaLysGlyGluAlaGlyProG 179

Qy      1664 AGAGCGCGGGTGGAGGT-----------GGGAGCAGG------CCACTGCCTCCAGCAC 1623
             ||   |||  |||               |||   |||     |||   ||||||   |
Db       179 lnGlyAlaArgGlySerGluGlyProGlnGlyValArgGlyGluProGlyProProGlyP 199

Qy      1622 CCACGTGTCCATTAGGG---AAGGGAGCTCCAGGCTTAGGG------CCTGGCAGGAAGC 1572
             ||           |||      |||   ||||||            ||||||  |||
Db       199 roAlaGlyAlaAlaGlyProAlaGlyAsnProGlyAlaAspGlyGlnProGlyAlaLys- 218

Qy      1571 TGGTCATCAGGCTGTCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGA 1512
               ::::::|||   |||                          ||||||   ||||||
Db       219 GlyAlaAsnGlyAlaPro--------------------------GlyIle---AlaGly 228

Qy      1511 ACACCTGCTTCTCCCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTG 1452
                |||    ||||||   ||||||                 :::|||    ||||||
Db       229 AlaProGlyPheProGlyAlaArgGly-----------ProSerGlyPro-GlnGlyPr 244

Qy      1451 AGAAGGTGAACCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGC 1392
                |||  ||||||    |||  |||   ||||||      |||        |||
Db       244 oSerGlyProProGlyProLysGlyAsnSerGlyGluProGly---AlaProGlyAsnLy 263

Qy      1391 ATGTGGCACCGGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGC 1332
                   ||||||:::      |||      ||||||
Db       263 sGlyAspThrGlyAlaLysGlyGluProGlyAlaThrGly------------------ 276

Qy      1331 CGAATCGCTGCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCA 1272
                 :::      |||  |||   |||            |||:::
Db       277 -------ValGlnGlyPro--ProGlyProAlaGlyGluGluGlyLysArgGlyAlaArg 293

Qy      1271 GGAACA--------------GCCCCAGGCTGCCCATCCGAACGC-------------- 1242
             |||                  ||||||    |||    ||||||
Db       294 GlyGluProGlyProSerGlyLeuProGlyProProGlyGluArgGlyGlyProGlySer 313

Qy      1241 -----------CTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCA 1194
                                |||:::||||||:::    |||:::
Db       314 ArgGlyPheProGlyAlaAspGlyValAlaGlyProLysGlyProSerGlyGluArgGly 333

Qy      1193 CGCCCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTG 1134
                 ||||||   |||         :::|||
Db       334 AlaProGlyProAlaGlyProLysGlySerPro------------------------- 344

Qy      1133 CCATCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGC 1074

Db       345 -------------------------------------------------------Gly 345

Qy      1073 ACAGCTGGTGCAGCCGGGGAAGCAGGGCGCCCAGGT-----------TCCGGAAAGCCA 1026
               ||||||      ||||||||||||    ||||||           :::|||   |||
Db       346 GluAlaGlyArgProGlyGluAlaGlyLeuProGlyAlaLysGlyLeuThrGlySerPro 365

Qy      1025 AGCGGGCCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTG 966
                ||||||   |||         |||   |||          |||
Db       366 GlySerProGlyProAsp------GlyLysThrGlyProProGlyProAlaGlyGlnAsp 383

Qy       965 GCTCGGTGGGGCCCAGCGCTGCCT----------------------------------- 942
               |||   |||||||||   |||
```

```
Db    384 GlyArgProGlyProAlaGlyProProGlyAlaArgGlyGlnAlaGlyValMetGlyPhe 403

Qy    941 CCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAA 882
          |||   |||                                |||      |||    ||||||:::
Db    404 ProGlyProLysGlyThrAla---------------GlyGluProGlyLysAlaGlyGlu 418

Qy    881 AGAGGCACTCCT------------------------------------------------ 870
          ||||||   |||
Db    419 ArgGlyLeuProGlyProProGlyAlaValGlyProAlaGlyLysAspGlyGluAlaGly 438

Qy    869 ---CCTGGGTGCCCAGGT---AGGGGGCCAGGGCACTGGTGTCCCAGTCAATGGCAGGCA 816
             |||  ||||||       ||||||                        |||       |||
Db    439 AlaGlnGlyAlaProGlyProAlaGlyProAlaGlyGluArgGlyGluGlnGlyProAla 458

Qy    815 GGA-----------------------GGTAGCCCAGGCAGCCCCCAAGACTGATCATGA 780
          |||                       |||     |||   |||    :::
Db    459 GlySerProGlyPheGlnGlyLeuProGlyProAlaGlyProProGlyGluAlaGlyLys 478

Qy    779 AGGCATAGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGTCCC------GGAAGAGGT 729
                    |||    |||||||||  |||   |||||||||         ||||||
Db    479 ProGlyGluGlnGlyValProGlyAspLeuGlyAlaProGlyProSerGlyAlaArgGly 498

Qy    728 CAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCA 669
          :::|||  |||   |||   |||               |||         ||||||   |||
Db    499 GluArg---GlyPheProGlyGluArgGlyValGlnGlyProProGlyProAlaGlyPro 517

Qy    668 CGCCCAGGATGAGCAGTGCCA----------------------------GCTCCAGGG 639
          |||              |||                             |||||||||
Db    518 ArgGlyAsnAsnGlyAlaProGlyAsnAspGlyAlaLysGlyAspThrGlyAlaProGly 537

Qy    638 GCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGC 579
          |||      ||||||:::                  :::   :::    |||||||||
Db    538 Ala------ProGlySerGlnGlyAlaProGlyLeuGlnGlyMetProGlyGluArgGly 555

Qy    578 TCAGCAGGATGCCCAAGGACAGTGCCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCAC 519
          :::||||||   |||          |||:::          |||
Db    556 AlaAlaGlyLeuProGly------ProLysGlyAspArgGlyAspAlaGlyProLysGly 573

Qy    518 GCCAGTGGTCACTGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAA 459
          |||   |||             ||||||        ||||||
Db    574 AlaAspGlySerProGlyLysAspGlyAlaArgGlyLeuThrGlyProIle--------- 590

Qy    458 TGCCCAGCACCATGGTCATGAACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCA 399
                                             ||||||   ||||||
Db    591 -----------------------------GlyProProGlyProAlaGlyAlaProGly 600

Qy    398 CATAGGTGATGCCTGCGGCCAAACACACCTCCAGGC------------------------ 363
             :::            |||:::   |||||||||
Db    601 AspLysGlyGluAlaGlyProSerGlyProProGlyProThrGlyAlaArgGlyAlaPro 620

Qy    362 ------CAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGT---GCCGCAGCA 312
                :::      ||||||  |||      ||||||   :::|||         |||
Db    621 GlyAspArgGlyGluAlaGlyProProGlyProAlaGlyPheAlaGlyProProGlyAla 640

Qy    311 GGCGGCTCACCCACAGCCTCT---------------------------------------- 291
            |||   |||   |||
Db    641 AspGlyGlnProGlyAlaLysGlyGluProGlyAspThrGlyValLysGlyAspAlaGly 660

Qy    290 ------GGACCATAGTGGGCCAGGCGGG---TAGGGCTCAGGGGGCCGTTCAGGCACTCC 240
          ||||||  :::||||||   |||       |||     |||     |||   ||
Db    661 ProProGlyPro-AlaGlyProAlaGlyProProGlyProIleGlyAsnValGlyAlaPr 680

Qy    239 AGAACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCA 180
          |                       |||      ||||||    ||||||     |||
Db    680 oGly-----------ProLysGlyProArgGlyAlaAlaGlyProPro-GlyAlaThr- 695

Qy    179 ACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGCTTC 120
                  |||        |||                |||       |||
Db    696 --------GlyPheProGlyAlaAlaGlyArgValGlyProProGlyProSerGlyAsnA 713

Qy    119 TCAGCCCATGCTCAACACCTGCTGCTGTGGGGCACCTCAGT--------------GGGG 75
             |||          |||      ||||||                            |||:
Db    713 laGlyProProGlyProProGlyProValGlyLysGluGlyGlyLysGlyProArgGlyG 733

Qy     74 ACACGTCTCATCACTCAGATCCTGGCCGA-----------------------------G 45
             ::|||              ||||||                              |
Db    733 luThrGlyProAlaGlyArgProGlyGluValGlyProProGlyProProGlyProAlaG 753

Qy     44 GCGCGCGGCTGTCACCCGGAGCC 22
            ||   :::   ||||||||||||
Db    753 lyGluLysGlySerProGlyAla 760
```

RESULT 44
US-08-931-820-1
; Sequence 1, Application US/08931820
; Patent No. 6010863
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Assay for collagen degradation
; NUMBER OF SEQUENCES: 4
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/931,820
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 96202596.1
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1057 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Collagen type I
US-08-931-820-1

Alignment Scores:

Pred. No.:	2.16e-18	Length:	1057
Score:	345.00	Matches:	249
Percent Similarity:	32.73%	Conservative:	42
Best Local Similarity:	28.01%	Mismatches:	308
Query Match:	5.54%	Indels:	291
DB:	3	Gaps:	51

US-09-759-143-110 (1-3410) x US-08-931-820-1 (1-1057)

```
Qy   2170 GGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCCCCAGCCCCCA 2111
            |||:::                |||                    ||||||   :|||   ||||||
Db     13 GlyIleSerValProGlyProMetGlyProSerGlyProArgGlyLeuProGlyProPro 32

Qy   2110 GCTGTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGC 2051
                                     ||||||      ||||||          |||
Db     33 GlyAlaPro---GlyProGlnGlyPheGlnGlyProProGlyGluProGlyGluProGly 51

Qy   2050 AGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAACAGGAGCGGGGAGCTG 1991
           ::::::           |||         |||           ::::::|||:::
Db     52 AlaSerGlyProMetGlyProArgGlyProProGlyProProGlyLysAsnGlyAspAsp 71

Qy   1990 GGA--------------------CCCAGTGAGGCAGGCCCTCCA--------------- 1967
           |||                    |||    |||   |||||||||
Db     72 GlyGluAlaGlyLysProGlyArgProGlyGluArgGlyProProGlyProGlnGlyAla 91

Qy   1966 ---------CCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTT 1916
                    |||   ||||||                      :::   |||
Db     92 ArgGlyLeuProGlyThrAlaGly---------------LeuProGlyMetLysGlyHis 106
```

```
Qy   1915 GTCAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCCTGCGGCAGACAC 1856
Db    107 ArgGlyPheSerGlyLeuAspGlyAlaLysGlyAspAlaGlyProAlaGlyProLysGly 126

Qy   1855 CATATAGGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCATAAACAGGGATGGGGC 1796
Db    127 GluProGlySerProGly-------------GluAsnGlyAlaProGlyGln--MetGlyP 142

Qy   1795 CA---------CCTGGGACAGCAGGAAGGCACTATCCAGGATGGCGAGGTCCAGGCAGAT 1745
Db    142 roArgGlyLeuProGlyGluArgGlyArg------ProGlyAlaProGlyProAlaGlyA 160

Qy   1744 GCCCCGGCCCGGAA--------------------CCACCCTGGCCTCGGTGGGCTCACC 1706
Db    160 laArgGlyAsnAspGlyAlaThrGlyAlaAlaGlyProProGlyProThrGlyProAlaG 180

Qy   1705 CACCACCACA---------------------CGTACGGAGACATCACAG---------- 1678
Db    180 lyProProGlyPheProGlyAlaValGlyAlaLysGlyGluAlaGlyProGlnGlyProA 200

Qy   1677 -----GCAGAGGCCCCGCAGAGCGCGGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCAC 1623
Db    200 rgGlySerGluGlyProGlnGlyValArgGly---------GluProGlyProProGlyP 217

Qy   1622 CCACGTGTCCATTAGGG---AAGGGAGCTCCAGGCTTAGGG------CCTGGCAGGAAGC 1572
Db    217 roAlaGlyAlaAlaGlyProAlaGlyAsnProGlyAlaAspGlyGlnProGlyAlaLys- 236

Qy   1571 TGGTCATCAGGCTGTCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGA 1512
Db    237 GlyAlaAsnGlyAlaPro--------------------------GlyIle---AlaGly 246

Qy   1511 ACACCTGCTTCTCCCGGTGGTAGAGGG---------------AGGCCAGTGTGTAGGGCA 1467
Db    247 AlaProGlyPheProGlyAlaArgGlyProSerGlyProGlnGlyProGlyGlyProThr 266

Qy   1466 GGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCA 1407
Db    267 GlyAlaArgGlyLeu-ValGlyGluProGlyProAlaGlySerLysGlyGluSerGlyAs 286

Qy   1406 CACTGTGGGACAGGCATGTGGCACCGGCAGCCACAGGGA-------------------- 1368
Db    286 nLys---GlyGluProGlySerAlaGly--ProGlnGlyProProGlyProSerGlyGlu 304

Qy   1367 AAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCTGCACCAGCCGGTCCA 1308
Db    305 GluGlyLysArgGlyProAsnGly------GluAlaGlySerAlaGlyProProGlyPro 322

Qy   1307 TGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACAGCCCCAGGCTGCCCATCC 1248
Db    323 ---ProGlyLeuArgGlyGly-----------------------ProGlySerArgGly 333

Qy   1247 GAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGCCCT 1188
Db    334 PheProGlyAlaAspGlyValAlaGlyProLysGlyProAlaGlyGluArgGlySerPro 353

Qy   1187 GGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCC 1128
Db    354 GlyProAlaGlyProLysGlySerPro-------------------------------- 362

Qy   1127 AGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACAGCT 1068
Db    363 ------------------------------------------------GlyGluAla 365

Qy   1067 GGTGCAGCCGGGGAAGCAGGGCGCCCAGGT------------TCCGGAAAGCCAAGCGGG 1020
Db    366 GlyArgProGlyGluAlaGlyLeuProGlyAlaLysGlyLeuThrGlySerProGlySer 385

Qy   1019 CCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGG 960
Db    386 ProGlyProAspGly------------LysThrGlyProPro------------------ 395

Qy    959 TGGGGCCCAGCG------------------CTGCCTCCTCAGCCACCA------------ 930
Db    396 ---GlyProAlaGlyGlnAspGlyArgProGlyProProGlyProProGlyAlaArgGly 414

Qy    929 ----------------------------------GCAGTGTGGCTGCTACGCAGGTGA 906
Db    415 GlnAlaGlyValMetGlyPheProGlyProLysGlyAlaAla------------------ 428

Qy    905 GGAAGATGAGGGTGAGCAGGCCAAAGAGGCACT------CCTCCTGGGTGCCCAGGT--- 855
Db    429 GlyGluProGlyLysAlaGlyGluArgGlyValProGlyProProGlyAlaValGlyPro 448

Qy    854 --------------------------------------------AGGGGGCCAGGGCAC 840
Db    449 AlaGlyLysAspGlyGluAlaGlyAlaGlnGlyProProGlyProAlaGlyProAlaGly 468

Qy    839 TGGTGTCCCAGTCAATGGCAGGCAGGA-------------------------GGTAGCCCA 804
Db    469 GluArgGlyGluGlnGlyProAlaGlySerProGlyPheGlnGlyLeuProGlyProAla 488

Qy    803 GGCAGCCCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGT--- 747
Db    489 GlyProProGlyGluAlaGlyLysProGlyGluGlnGlyValProGlyAspLeuGlyAla 508

Qy    746 CCGGGTCCC------GGAAGAGGTCAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCT 693
Db    509 ProGlyProSerGlyAlaArgGlyGluArg---GlyPheProGlyGluArgGlyValGln 527

Qy    692 GGCCACAGAAGTCCAGCAGCCCCACGCCCAGGATGAGCAGTGCCA--------------- 648
Db    528 GlyProProGlyProAlaGlyProArgGlyAlaAsnGlyAlaProGlyAsnAspGlyAla 547

Qy    647 ---------------GCTCCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGC 603
Db    548 LysGlyAspAlaGlyAlaProGlyAla------ProGlySerGlnGlyAlaProGlyLeu 565

Qy    602 CGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCAAGGACAGTGCCCAGATGAAGG 543
Db    566 GlnGlyMetProGlyGluArgGlyAlaAlaGlyLeuProGly------ProLysGlyAsp 583

Qy    542 GCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTAGGAGCG---GGA 486
Db    584 ArgGlyAspAlaGlyProLysGlyAlaAspGly---------SerProGlyLysAspGly 600

Qy    485 CACAGA-----------------------CCAGGCCCAGCACTGGACCAATGCCCA--- 453
Db    601 ValArgGlyLeuThrGlyProIleGlyProProGlyProAla------GlyAlaProGly 618

Qy    452 --------------------------------------------------GCACCA 447
Db    619 AspLysGlyGluSerGlyProSerGlyProAlaGlyProThrGlyAlaArgGlyAlaPro 638

Qy    446 TGGTCATGAACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGC 387
Db    639 GlyAspArgGlyGluProGlyProProGlyProAlaGlyPheAlaGly------------ 654

Qy    386 CTGCGGCCAAACACACCTCCAGGC------------CAAAGGTTAGCAGGTTGACCAGCA 339
Db    655 ---------------ProProGlyAlaAspGlyGlnProGlyAlaLysGlyGluProGly 669

Qy    338 AGAGCTGGGCTTTCCGGTGCCGCAGCAGGCGGCTCACCCACAGCCTCTGGACCATAGTGG 279
Db    670 AspAlaGlyAlaLysGlyAspAla------GlyProProGlyProAlaGlyPro-AlaGl 687

Qy    278 GCCA------------GGCGGGTAGGGCTCAGGGGGCCGTTCAGGCACTCCAGAACTGCT 231
Db    687 yProProGlyProIleGlyAsnValGlyAlaProGlyAlaLysGlyAla----------- 703

Qy    230 TCGTCTCGGCTCTGCT------CCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAACT 177
Db    704 ----ArgGlySerAlaGlyProProGlyAlaThrGlyPheProGlyAlaAlaGlyArgVa 722

Qy    176 GCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGCTTCTCA 117
```

```
Db        722 lGlyProProGlyProSerGlyAsnAlaGlyProProGlyProProGlyProAlaGlyLy 742

Qy        116 GCCCATGCTCAACACCTGCTGCTGTGGG------GCACCTCAGTGGGGACACGTCTCATC 63
              :                           |||        |||:::  |||:::|||
Db        742 sGlu--------------------GlyGlyLysGlyProArg-GlyGluThrGlyProA 755

Qy         62 ACTCAGATCCTGGCCGA-----------------------------GGCGCGCGGCTGT 33
                        ||||||                                |||   :::    |
Db        755 laGlyArgProGlyGluValGlyProProGlyProProGlyProAlaGlyGluLysGlyS 775

Qy         32 CACCCGGAGCC 22
              |||||||||||
Db        775 erProGlyAla 778
```

RESULT 45
US-08-642-255-73
; Sequence 73, Application US/08642255
; Patent No. 5773249
; GENERAL INFORMATION:
; APPLICANT: CAPPELLO, Joseph
; APPLICANT: FERRARI, Franco A.
; TITLE OF INVENTION: High Molecular Weight Collagen-Like
; TITLE OF INVENTION: Protein Polymers
; NUMBER OF SEQUENCES: 135
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/642,255
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ROWLAND, Bertram I.
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A55556-3/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 633 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-642-255-73

Alignment Scores:

Pred. No.:	1.97e-18	Length:	633
Score:	344.50	Matches:	245
Percent Similarity:	33.54%	Conservative:	32
Best Local Similarity:	29.66%	Mismatches:	289
Query Match:	5.53%	Indels:	261
DB:	1	Gaps:	54

US-09-759-143-110 (1-3410) x US-08-642-255-73 (1-633)

```
Qy       2314 AAAAACCCTTCTCTAGGTGTGTCTCAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGT 2255
              :::||||||       ||||||:::|||                  |||
Db         12 GluAsnPro------GlyValThrGln--------------LeuAsnArgLeuAlaAla 24

Qy       2254 AATCCACCTGCAGAGTCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAA 2195
              :::||||||      |||              |||      |||:::|||   |||
Db         25 HisProProPheAlaSerAspProMetGlyAlaHisGlyProAlaGly---ProLysGly 43

Qy       2194 GTCCAGACTGAAACCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAG 2135
                                 :::|||   |||               |||         |||
Db         44 AlaHisGlyProAlaGlyProLysGlyAlaGln------GlyProAlaGlyProGlyGly 61

Qy       2134 AGGAGAGGGACGCCCCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGTGGC 2075
                ||||||   |||   ||||||      |||                |||    ||||||
Db         62 SerArgGlyAspProGlyProProGlyAlaGln------GlyProAlaGlyProGlyGly 79

Qy       2074 AGCAGA---GAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGC-- 2020
              ||||||   :::|||          ||| |||               |||    ||||||
Db         80 SerArgGlyAspProGlyProProGly-AlaGlnGlyProAlaGlyProGlyGlySerAr 99

Qy       2019 ------------CCCATGGGGCTAACAGGA---GCGGGGAGCTGGGACCCAGTGAGGCA 1976
                          |||   |||          |||  ||| |||   |||        |||
Db         99 gGlyAspProGlyProProGlyAlaGlnGlyProAla-GlyProGlyGlySerArgGlyA 119

Qy       1975 GGCCCTCCACCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTT 1916
                |||      |||
Db        119 spProGlyProProGlyAlaGlnGly------------------------------- 127

Qy       1915 GTCAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCCTGCGGCAGACAC 1856
                          |||                  |||   |||||||||
Db        128 -----------ProAlaGlyProGlyGlySerArgGlyAspProGly------------ 139

Qy       1855 CATATAGGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCATAAACAGGGATGGGGC 1796
                                   |||      |||         |||         |||
Db        140 -------------ProProGlyAlaGlnGlyProAlaGlyProGlyGlySerArgGlyA 155

Qy       1795 CACCTGGGACAGCAGGAAGGCACTATCCAGGATGGCGAGGTCCAGGCAGATGCCCCGGCC 1736
               ||||||     |||   |||    |||           ||||||          |||
Db        155 spProGlyProProGlyAlaHisGlyProAla------GlyProLysGlyAlaHisGly- 172

Qy       1735 CGGAACCACCCTGGCCTCGGTGGGCTCACCCACCACCACACGTACGGAGACATCACAG-- 1678
                    |||     |||:::   ||||||   |||          :::      |||
Db        173 -----ProAlaGlyProLysGlyAlaHisGlyProAlaGlyProLysGlyAlaHisGlyP 191

Qy       1677 ----------------GCAGAGGCCCCGCAGAGCGCGGGTGGAGGTGGGAGCAGG---- 1639
                              |||:::   |||        ||||||   |||||||||||
Db        191 roAlaGlyProLysGlyAlaGlnGlyPro------AlaGlyProGlyGlySerArgGlyA 209

Qy       1638 --CCACTGCCTCCAGCACCCACGTGTCCATTAGGG-----------AAGGGAGCTCCAG 1593
                |||   ||||||            |||   |||              :::|||  ||||
Db        209 spProGlyProProGlyAlaGlnGlyProAlaGlyProGlyGlySerArgGlyAspProG 229

Qy       1592 GCTTAGGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCTAGCACCTCCAGTGT 1533
              ||      ||||||   :::                                    |||
Db        229 ly---ProProGlyAlaGln---------------------------GlyProAlaG 238

Qy       1532 CCCCTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGTAGAGGGAGGCCAGTGTGT 1473
                ||| |||        |||   |||      ||||||   :::|||   |||
Db        238 lyPro-GlyGlySerArgGlyAspProGlyProProGlyAlaGlnGly---ProAlaGly 256

Qy       1472 AGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAGCTGTCACCA 1413
                 ||||||   |||                                        |||
Db        257 ProGlyGlySerArgGly---------------------------------AspPro 264

Qy       1412 CGGCCACACTGTGGGACAGGCATGTGGCACCGGCAGCCACAGGGAAAGCTGCCACACTGG 1353
                |||      |||
Db        265 GlyProProGlyAlaGlnGly------------------------------------- 271

Qy       1352 CCAAATAGACTGCTCGAGTGCCGAATCGCTGCACCAGCCGGTCCATGACCAGAGAGA--- 1296
                                                 ||||||||||||      |||
Db        272 ---------------------------------ProAlaGlyProGlyGlySerArgGly 280

Qy       1295 ---------AGACCAGGGAGATGGCGCACTGCAGGAACAGCC--------------CCA 1260
                        ||||||              ||||||                     |||
Db        281 AspProGlyProProGlyAlaGlnGlyProAlaGlyProGlyGlySerArgGlyAspPro 300

Qy       1259 GGCTGCCCATCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTC 1200
              |||   |||                ||||      :::||||||:::      |||
Db        301 GlyProProGlyAla---------HisGlyProAlaGlyProLys---GlyAlaHisGly 316
```

```
Qy       1199 TGGGCACGCCCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCA 1140
                 |||  |||    |||     |||
Db        317 ProAlaGlyProLysGlyAlaHisGlyPro------------------------------ 326

Qy       1139 TGAGTGCCATCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGTGCGGGGCATGC 1080
                                    |||   |||:::  |||           |||
Db        327 --------------------AlaGlyProLysGlyAlaHisGlyProAlaGlyProLys 339

Qy       1079 GGCAGC---------------ACAGCTGGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCC 1035
              |||:::                |||:::  |||:::   |||   ||||||:::
Db        340 GlyAlaGlnGlyProAlaGlyProGlyGlySerArgGlyAspProGlyProProGlyAla 359

Qy       1034 GGAAAGCCAAGCGGGCCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCC 975
                    |||:::|||||||||     |||   |||            ||||||
Db        360 GlnGlyProAlaGlyProGly---GlySerArgGlyAspPro------GlyProPro--- 375

Qy        974 CTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCT-----------------CCTCAGCCA 933
                                |||||||||   |||                |||   |||
Db        376 ---------GlyAlaGlnGlyProAlaGlyProGlyGlySerArgGlyAspProGlyPro 392

Qy        932 CCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGCACT 873
              |||                           |||    |||     |||   ||||||
Db        393 ProGlyAlaGln---------------GlyProAlaGlyProGlyGlySerArgGlyAsp 407

Qy        872 ------CCTCCTGGGTGCCCAGGT---AGGGGGCCAGGGCACTGGTGTCCCAGTCAATGG 822
                    |||||||||     |||    ||||||
Db        408 ProGlyProProGlyAlaGlnGlyProAlaGlyPro------------------------ 419

Qy        821 CAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCATAGACAGAGTAGG 762
                    |||||||||   |||   |||
Db        420 ------GlyGlySerArgGlyAspPro-------------------------GlyPro 428

Qy        761 CCTGGCGACAGTGGTCCG---GGTCCCGGAAGAGGTCAGAGAGCAGGGCCTCCAGTGGAG 705
              ||||||      ||||||   |||||||||   |||   |||       |||
Db        429 ProGlyAlaGlnGlyProAlaGlyProGly---GlySerArgGlyAspPro--------- 444

Qy        704 TGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGCCCAGGATGAGCAGTGCCAGCT 645
                          ||||||         ||||||   |||:::
Db        445 ---GlyProProGlyAlaHisGlyProAlaGlyProLys-------------------- 456

Qy        644 CCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAA 585
                 ||||||   |||                    :::   |||   |||
Db        457 ---GlyAlaHisGlyPro--------------------AlaGlyProLysGly---Ala 467

Qy        584 AGAGGCTCAGCAGGATGCCCAAGGACAGTGCCCAGATGAAGGGCCGGCGGCGGCCATAGC 525
                 |||   ||||||                |||:::      |||
Db        468 HisGlyProAlaGly---------------ProLysGlyAlaHisGlyProAlaGlyPro 482

Qy        524 GTCCACGCCAGTGGTCACTGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTG 465
                    |||   |||                 |||:::         |||||||||
Db        483 LysGlyAlaGlnGlyProAlaGlyProGlyGlySerArgGlyAspProGlyProProGly 502

Qy        464 GACCAATGCCCAGCACCATGGTCATGAACTTCTCCTCTACCCCCACTTCCAGCAGCAGAG 405
                 |||   ||||||            :::         |||      |||   |||:::
Db        503 AlaGlnGlyProAlaGlyProGlyGlySerArgGlyAspProGlyProProGlyAlaGln 522

Qy        404 GCGGCACATAGGTGATGCCTGCGGCCAAACACACCTCCAGGCCAAAGGTTAGCAGGTTGA 345
                                      |||   |||   |||   |||      |||
Db        523 Gly---------------------ProAlaGlyProGlyGlySerArg------GlyAsp 533

Qy        344 CCAGCAAGAGCTGGGCTTTCCGGT---GCCGCAGCAGGCGGCTCA--------------- 303
              |||         |||   |||   |||      |||||||||
Db        534 ProGlyProProGlyAlaGlnGlyProAlaGlyProGlyGlySerArgGlyAspProGly 553

Qy        302 ---CCCACAGCCTCTGGACCATAGTGGGCCAGGCGGGTAGGGCTCAGGGGGCCGTTCAGG 246
                 |||   |||   |||||| :::|||||||||       ||||||   |||     ||
Db        554 ProProGlyAlaGlnGlyPro-AlaGlyProGly------GlySerArgGlyAspProGl 571

Qy        245 CACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTG 186
              |  |||                                :::    |||   |||
Db        571 yProPro--------------------GlyAlaGlnGlyProAlaGlyProGlyGlySe 584

Qy        185 CCGCCAACTGCCTAGGAATCAGCCAGGCGCCCATTTC---TGCCAGCCCTTTGGTGCC-- 131
                 |||      |||        |||||||||||          |||   ||||||
Db        584 rArgGlyAspPro---GlyProProGlyAlaHisGlyProAlaGlyProLysGlyAlaHi 603

Qy        130 -GGTCCAGCTTCTCAGCCCATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTGGGGACA 72
               |||||||||                                              |||
Db        603 sGlyProAla---------------------------------------------GlyPr 608

Qy         71 CGTCTCATCACTCAGATCCTGGCCGAGGCGCGCGGCTGTCACCCGGAGCCAGCGCGTGCA 12
                       :::::: ||||||||||||         :::||||||   |||          :
Db        608 oLysGlyAlaMet-AspProGlyArg---TyrGlnLeuSerAlaGlyArgTyrHisTyrG 627

Qy         11 GGCTGGTT 4
              ::||||||
Db        627 lnLeuVal 629
```

Search completed: June 9, 2003, 22:35:21
Job time : 194.5 secs

```
US-09-759-143-110 (1-3410) x Q9FVL6 (1-604)

Qy     193 GAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAGTGCCTGAAC 252
           |||  |||          |||||        |||  :::          |||
Db      25 GluSerArgPheAlaGlnLeuGluIleArgSerAspSerSerSerProArgValSerAsn 44

Qy     253 GGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCAGAGGCTGTGGGTGAGCCGCCT 312
           |||  ::::::  :::  |||                  :::       ||||||
Db      45 GlyGluMetAsnAspSerAsnLeu------------------ProLeuProProProPro 58

Qy     313 GCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGT 372
           :::|||:::             |||:::|||       :::       |||::::::
Db      59 -ValArgAsnSerLeuLeuThrLeuIleLeuSerCysThrValAlaAlaGlyValGlnPh 78

Qy     373 GTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGA 432
              |||     ::::::  :::  |||  :::      :::|||:::|||
Db      78 eGlyTrpAlaLeuGlnLeuSerLeuLeuThrProTyrIleGlnThrLeuGlyIleGluHi 98

Qy     433 GAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCT 492
              |||  :::  :::           ||||||:::  |||||||||     |||
Db      98 sAlaPheSerSerPheIleTrpLeuCysGlyProIleThrGlyLeuValValGlnProCy 118

Qy     493 CCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTG 552
           :::|||     ||||||        :::||||||||||||||||||||||||||||::
Db     118 sValGlyIleTrpSerAspLysCysHisSerLysTyrGlyArgArgArgProPheIlePh 138

Qy     553 GGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGC 612
           :      ::::::  |||  :::::::::   :::|||         |||:::  ||
Db     138 eIleGlyAlaValMetIleSerIleAlaValIleIleIle---------GlyPheSerAl 155

Qy     613 A------GGGCTGCTGTGCCCGGAT-----------------------------CCCAG 636
           |      |||  |||   |||                                 ||
Db     155 aAspIleGlyTyrLeuLeuGlyAspThrLysGluHisCysSerThrPheLysGlyThrAr 175

Qy     637 GCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGT 696
           |                :::  ::::::|||    :::||||||
Db     175 gSerArgAlaAlaIleValPheValValGlyPheTrpMetLeuAspLeuAlaAsnAsnTh 195

Qy     697 GTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCG 756
                 |||      ||||||||:::||||||           ||||||
Db     195 rValGlnGlyProAlaArgAlaLeuLeuAlaAspLeu---SerGlyProAspGlnArgAs 214

Qy     757 CCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCT 816
              |||   :::|||:::      ::::::::|||    ||||||:::
Db     214 nThrAlaAsnAlaValPheCysSerTrpMetAlaValGlyAsnIleLeuGlyPheSerAl 234

Qy     817 GCCTGCCATTGAC------------TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCAC 864
              |||                 |||                  |||:::|||
Db     234 aGlyAlaSerGlyGlyTrpHisArgTrp---------------PheProPheLeuThrAs 249

Qy     865 CCAGGAGGAGTGC---------------------CTCTTTGGCCTGCTCACCCTCATCTT 903
            :::      |||                      |||:::   ::::::||
Db     249 nArgAlaCysCysGluProCysGlyAsnLeuLysAlaAlaPheLeuValAlaValValPh 269

Qy     904 CCTCACC---TGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCAC 960
           |||||||   |||          ||||||     |||   |||      |||   |||
Db     269 eLeuThrLeuCysThrLeuValThrLeuTyrPheAlaAsnGluValProLeuSerProLy 289

Qy     961 CGAG---CCAGCAGAAGGGCTGTCGGCCCCCTCCTTG---TCGCCCCAC----------- 1003
            :::                   |||||||||  |||   ||||||
Db     289 sGlnTyrLysArgMetSerAspSerAlaProLeuLeuAspSerProGlnAsnThrGlyPh 309

Qy    1004 ----TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTG----------------- 1042
                      :   :::        |||    :::||||||
Db     309 eAspLeuSerGlnSerLysArgGluLeuGlnTyrArgAsnSerValAlaAsnAsnGluSe 329

Qy    1042 ------------------------------------------------------------ 1042

Db     329 rGluMetGlyHisValAlaAspAsnSerProLysAsnGluGluGlnArgProAspLysAs 349

Qy    1043 ---------------------------GGCGCCCTGCTTCCCCGGCTGCACCAGCTGTG 1074
                                      ||||||:::|||       |||
Db     349 pGlnGlyAspSerPheAlaAspSerProGlyAlaValLeuValAsnLeuLeuThrSerLe 369

Qy    1075 CTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGC 1134
                 :::|||    :::     :::    :::           :::|||:::::
Db     369 uArgHisLeuProProAlaMetHisSerValLeuIleValMetAlaLeuThrTrpLeuSe 389

Qy    1135 ACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGT 1194
           :            |||   ||||||   ||||||::::::|||       :::|||   |||
Db     389 rTrpPheProPhePheLeuPheAspThrAspTrpMetGlyArgGluValTyrHisGlyAs 409

Qy    1195 GCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGG 1254
            |||:::   |||                :::            |||:::::|||||||||   ||
Db     409 pProLysGlyGlu------AlaAspGluValAsnAlaTyrAsnGlnGlyValArgGluGl 427

Qy    1255 CAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCG 1314
            |:::  ||||||  |||               :::        |||   |||  :::::::::
Db     427 yAlaPheGlyLeuLeuLeuAsnSerValValLeuGlyValSerSerPheLeuIleGluPr 447

Qy    1315 GCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGT----------------- 1357
            :::   :::         |||:::|||   |||:::      |||
Db     447 oMetCysLysTrpIleGlySerArgLeuValTrpAlaValSerAsnPheIleValPheVa 467

Qy    1358 ---------------------GTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCT 1395
                               ::::::          ::::::|||               ::
Db     467 lCysMetAlaCysThrAlaIleIleSerValValSerIleSerAlaAsnThrGlnGlyVa 487

Qy    1396 GTCCCACAGTGTGGCCGTGGTG------ACAGCTTCAGCCGCCCTCACCGGGTTCACCTT 1449
           :   |||      :::                  |||         |||||||||   |||:::
Db     487 lGlnHisValIleGlyAlaThrArgSerThrGlnIleAlaAlaLeuValValPheSerLe 507

Qy    1450 CTCAGCCCTGCAGATC---------------CTGCCCTACACACTGGCCTCCCTCTACCA 1494
                 :::   :::                  :::|||:::::::::    :::
Db     507 uLeuGlyIleProLeuAlaValThrTyrSerValProPheSerIleThrAla-------- 524

Qy    1495 CCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGA 1554
                                         :::       |||   |||
Db     525 ------------------------GluLeuThrAlaAspAlaGly-------------- 531

Qy    1555 GGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGG 1614

Db     531 ------------------------------------------------------------ 531

Qy    1615 ACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTC 1674
                       ||||||   ||||||
Db     532 -------------GlyGlyGlnGlyLeu-------------------------------- 536

Qy    1675 TGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCC 1734
                                       :::|||            |||   ||||||||
Db     537 ------------------------AlaIleGlyValLeuAsnLeuAlaIleValValPr 548

Qy    1735 GGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGT 1794
           |          :::  |||                |||:::   |||              ::
Db     548 oGlnMetValValSerLeuGlyAlaGlyProTrpAspAlaLeuPheGlyGlyGlyAsnIl 568

Qy    1795 GGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATAT 1854
           :   |||:::      :::   |||:::        |||:::      :::   |||   ::
Db     568 e---ProAlaPheAlaLeuAlaSerLeuAlaAlaLeuAlaAlaGlyIlePheAlaMetLe 587

Qy    1855 G 1855
           :
Db     587 u 587


Search completed: June  9, 2003, 22:32:18
Job time : 317.5 secs
```

```
Qy      268 ACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGTGAGCCGCCT-----------GCT 315
            |||:::||||||       |||  :::         ||||||              |||
Db        3 ThrGlnLeuAlaThrLysAsnProSerSerGlnProProProProProProGluLeuAla 22

Qy      316 GCGGCACCGGAAAGCCCAGCTC---------TTGCTGGTCAACCTGCTAACCTTTGGCCT 366
            |||  |||   |||   |||         :::::::|||      :::       |||::
Db       23 AlaProProSerSer-AsnLeuLysLysIleValValValAlaSerIleAlaAlaGlyVa 42

Qy      367 GGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGT 426
             ::::          |||     ::::::  :::   |||    :::      :::|||::
Db       42 lGlnPheGlyTrpAlaLeuGlnLeuSerLeuLeuThrProTyrValGlnLeuLeuGlyIl 62

Qy      427 AGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGT 486
            :          :::          :::          ||||||:::   |||::::::
Db       62 eProHisThrTrpAlaAlaPheIleTrpLeuCysGlyProIleSerGlyMetLeuValGl 82

Qy      487 CCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTT 546
             |||:::::|||      ||||||         |||:::||||||||||||||||||
Db       82 nProIleValGlyTyrHisSerAspArgCysThrSerArgPheGlyArgArgArgProPh 102

Qy      547 CATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCC----- 601
            ||||  |||   :::       :::   ::::::::|||||||||      |||
Db      102 eIleAlaAlaGlyAlaAlaSerValAlaValAlaValPheLeuIleGlyTyrAlaAlaAs 122

Qy      602 ----GGCTGGCTAGCAGGG---CTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACT 654
                |||   ::::::|||       :::         |||:::        :::|||::
Db      122 pLeuGlyHisMetSerGlyAspSerIleGlyLysSerProLysIleArgAlaIleAlaIl 142

Qy      655 GCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGA 714
            :   ::::::|||       :::||||||            :::         |||
Db      142 ePheValValGlyPheTrpIleLeuAspValAlaAsnAsnMetLeuGlnGlyProCysAr 162

Qy      715 GGCCCTGCTCTCTGACCTCTTC---CGGGACCCGGACCACTGTCGCCAGGCCTACTCTGT 771
             ||||||||:::||||||                   |||   |||     :::::
Db      162 gAlaLeuLeuAlaAspLeuSerGlyThrSerGlnLysLysThrArgLeuAlaAsnAlaLe 182

Qy      772 CTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC-- 829
             :::::::|||   :::::::::|||      |||||||||          |||
Db      182 uPheSerPhePheMetAlaValGlyAsnIleLeuGlyTyrAlaAlaGlyAlaTyrThrHi 202

Qy      830 ---------------------TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCA 867
                                   |||   |||         |||
Db      202 sLeuTyrLysIleAlaProPheThrLeuThrLysAlaCysAspValTyrCysAlaAsnLe 222

Qy      868 GGAGGAGTGCCTCTTT---GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCAC 924
             :::   |||   |||       :::       |||:::    ||||||    :::|||   ||
Db      222 uLysThrCysPhePheIleSerIleAlaLeuLeuLeuValLeuThrValIleAlaLeuTh 242

Qy      925 ACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTC--- 981
            |   :::             :::|||   ::::::|||:::|||        ::::::
Db      242 rTyrVal------LysGlu-LysGlnTrpSerPro-GluIleAlaLysThrValAlaGly 259

Qy      982 --------------------GGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGC 1020
                                |||||||:::          |||||||||
Db      260 ValAspGlyAspAspGluAspGlyProValGluAlaAlaProLeu--------------- 274

Qy     1021 CCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCG 1080
                          |||       |||   :::        |||      |||
Db      275 -----------Pro-PhePheGlyGluIlePheAlaAlaLeuLysAsnLeu-------- 287

Qy     1081 CATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCAT 1140
                |||   :::|||   |||    :::              :::|||:::|||
Db      288 ----GlnArgProMetArgIleLeuLeuLeuValThrCysLeuAsnTrpIleAlaTrpPh 306

Qy     1141 GACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAG 1200
              |||   ||||||    ||||||:::::::|||             :::|||   |||
Db      306 eProPheLeuLeuPheAspThrAspTrpMetGlyArgGluValTyrGlyGlySerSerAs 326

Qy     1201 AGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCT 1260
             :::                          :::   ||||||   |||||||||   |||:::||
Db      326 pSerSerAlaThrAlaGlyGluSerLysLeuTyrAspArgGlyValArgAlaGlyAlaLe 346

Qy     1261 GGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT 1320
             |||||||    |||          :::              ||||||    ::::::    :::
Db      346 uGlyLeuMetLeuAsnSerValValLeuGlyPheThrSerLeuGlyValGluLeuMetAl 366

Qy     1321 GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGC 1380
               ::::::   |||   :::             |||         :::       |||   :::
Db      366 aArgLysLeuGly---LysValLysArgLeuTrpGlyIleValAsnPheLeuLeu----- 383

Qy     1381 CGGTGCCACATGCCTGTCC----------------------------------------- 1399
                |||   ||||||:::
Db      384 ----AlaIleCysLeuAlaLeuThrIleLeuIleThrLysLeuAlaGluSerGlyArgHi 402

Qy     1400 -------------CACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCAC 1446
                         :::   |||               |||   |||         :::|||
Db      402 sTyrProThrValAsnGlyValValSerThrThrProSerProProSerSerGlyIleLy 422

Qy     1447 CTTCTCAGCCCTGCAGATC---------------------------------------CT 1467
                       ||||||   :::                                      ::
Db      422 sAlaGlyAlaLeuThrLeuPheAlaValMetGlyIleProGlnAlaIleThrTyrSerVa 442

Qy     1468 GCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAG---------------------- 1504
             :|||:::   |||||||||::::::        :::
Db      442 lProPheAlaLeuAlaSerIlePheSerAsnAspSerGlyAlaGlyGlnGlyLeuSerLe 462

Qy     1505 ----------------CAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 1548
                          :::   :::|||:::
Db      462 uGlyValLeuAsnLeuSerIleValIleProGln------------------------- 473

Qy     1549 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGA----------- 1597
                             ::::::      |||      ||||||
Db      474 -------------MetLeuValSerValAlaSerGlyProPheAspAlaLeuPheGlyGl 489

Qy     1598 ----------GCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTG----- 1642
                            |||       |||     |||     |||         ||||||:::
Db      489 yGlyAsnLeuProAlaPheValValGlyAlaValAlaAlaAlaAlaSerGlyValLeuAl 509

Qy     1643 ----------CTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTC 1687
                        ||||||    ||||||                |||   |||      |||
Db      509 aLeuThrTyrLeuProSerProProProAspLeuProAlaLysAlaLeuThrVal 527
```

RESULT 45

```
Q9FVL6
ID   Q9FVL6      PRELIMINARY;      PRT;   604 AA.
AC   Q9FVL6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Sucrose transporter-like protein.
GN   SUT2.
OS   Lycopersicon esculentum (Tomato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4081;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20359188; PubMed=10899981;
RA   Barker L., Kuhn C., Weise A., Schulz A., Gebhardt C., Hirner B.,
RA   Hellmann H., Schulze W., Ward J.M., Frommer W.B.;
RT   "SUT2, a putative sucrose sensor in sieve elements.";
RL   Plant Cell 12:1153-1164(2000).
DR   EMBL; AF166498; AAG12987.1; -.
DR   InterPro; IPR003662; sub_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   TIGRFAMS; TIGR01301; GPH_sucrose; 1.
SQ   SEQUENCE   604 AA;  65082 MW;  3403349B1692B2AF CRC64;

Alignment Scores:
Pred. No.:                5.05e-13     Length:          604
Score:                    327.50       Matches:         141
Percent Similarity:       37.29%       Conservative:    98
Best Local Similarity:    22.00%       Mismatches:      238
Query Match:              5.10%        Indels:          164
DB:                       10           Gaps:            21
```

```
Db         26 LeuPheLeuAlaCysMetValAlaGlyGlyValGlnTyrGlyTrpAlaLeuGlnLeuSer 45

Qy        395 TATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTG 454
                 :::   |||   :::       :::|||:::                 :::::::::
Db         46 LeuLeuThrProTyrValGlnThrLeuGlyIleProHisAlaLeuThrSerValMetTrp 65

Qy        455 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 514
                    ||||||:::  ||||||:::     |||   :::|||    ||||||
Db         66 LeuCysGlyProIleAlaGlyLeuIleValGlnProCysValGlyLeuTyrSerAspLys 85

Qy        515 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 574
                             ||||||||||||||||||||          :::   |||
Db         86 CysThrSerSerLeuGlyArgArgArgProPheIleLeuThrGlyCysIleIleIleCys 105

Qy        575 CTGAGCCTCTTTCTCATC---------CCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGC 625
              :::|||:::   :::|||                  |||:::      |||
Db        106 IleSerValIleValIleGlyPheSerSerAspIleGlyTyrAlaLeuGlyAspThrThr 125

Qy        626 CCGGAT------------------CCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGC 667
                 |||                  ||||||          |||      |||||||||
Db        126 GluAspCysLysValTyrArgGlyProArgTyrHisAlaAlaAlaAlaPheIleLeuGly 145

Qy        668 GTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCT 727
                    ||||||||||||                  |||      ||||||::::::
Db        146 PheTrpLeuLeuAspPheSerAsnAsnThrValGlnGlyProAlaArgAlaLeuMetAla 165

Qy        728 GACCTCTTC------CGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTC 781
              ||||||             |||                |||   :::::::::
Db        166 AspLeuSerGlyArgHisGlyProSer---------AlaAlaAsnAlaIlePheCysSer 182

Qy        782 ATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATT---GACTGGGACACC 838
                 ::::::||||||      |||||||||      :::      ||||||
Db        183 TrpMetAlaLeuGlyAsnIleLeuGlyTyrSerSerGlySerThrAsnAspTrp------ 200

Qy        839 AGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGC--------------------- 877
                          |||:::|||   |||:::      |||
Db        201 HisLysTrpPheProPheLeuMetThrArgAlaCysCysGluAlaCysAlaAsnLeuLys 220

Qy        878 CTCTTTGGCCTGCTCACCCTCATCTTCCTC---ACCTGCGTAGCAGCCACACTGCTGGTG 934
                       |||:::   ::::::||||||         |||   |||::::::
Db        221 AlaAlaPheLeuValAlaValValPheLeuGlyLeuSerThrAlaValThrMetValPhe 240

Qy        935 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 994
              |||   |||   ||||||   |||          |||:::
Db        241 AlaArgGluValAlaLeuAspProValAlaAlaAlaLys--------------------- 253

Qy        995 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGC--------- 1045
                                                 ||||||   |||
Db        254 ------------------------------------ArgAsnGluGlyGluAlaSer 260

Qy       1046 ---------GCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTG 1096
                       |||:::         :::      |||         ||||||      |||
Db        261 GlyLeuLeuAlaValPheLysGlyMetLysAsnLeuProValGlyMetProSerValLeu 280

Qy       1097 CGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTT 1156
                          :::         :::|||::::::         |||   ||||||
Db        281 ------------IleValThrGlyLeuThrTrpLeuSerTrpPheProPheIleLeuPhe 296

Qy       1157 TACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACC 1216
                 ||||||::::::|||        :::|||   |||   |||         |||
Db        297 AspThrAspTrpMetGlyArgGluIleTyrHisGlyArgProAspGlySerProAlaGlu 316

Qy       1217 GAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAG 1276
                          :::   ||||||||||||   |||:::   ||||||   |||
Db        317 ValThrAla------PheGlnGluGlyValArgGlnGlyAlaPheGlyLeuLeuLeuAsn 334

Qy       1277 TGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACT 1336
                    :::      :::   |||   :::::::::   :::   :::|||   |||
Db        335 SerIleValLeuGlyIleSerSerPheLeuIleGluProMetCysArgArgLeuGlyAla 354

Qy       1337 CGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTG 1396
              |||||||||::::::   |||   |||         ||||||      |||
Db        355 ArgAlaValTrpValMetSerSerAlaValValCysValAlaMetAlaAla--------- 371

Qy       1397 TCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCC 1456
                       |||:::|||::::::|||
Db        372 ---------ValSerValLeuSerAla--------------------------------- 377

Qy       1457 CTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTG 1516
                             ::::::|||
Db        378 ---------------TrpSerLeu------------------------------------ 380

Qy       1517 CCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTC 1576
                          ||||||   ||||||   |||   :::|||:::
Db        381 ------------GlyAspPheGlyGly---SerValGlnAspAlaAlaArg--------- 392

Qy       1577 CTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGT 1636
                                  ||||||      |||   |||   |||
Db        393 ---------------------AlaProAlaGluGluGlyGlyValArgAlaSerAlaLeu 405

Qy       1637 GGCCTGCTCCCA------------CCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGAT 1684
                 |||                  |||         ||||||
Db        406 AlaLeuPheValPheLeuGlyLeuProPheAlaValLeuCys------------------ 419

Qy       1685 GTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGC 1744
                 ||||||         |||               |||         |||:::|||
Db        420 ---SerValProPheAlaValThrAlaGlnLeuAlaAlaSerArgGlyGlyGlyGlnGly 438

Qy       1745 ATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCC 1804
              :::|||               :::|||:::   :::   ::::::   |||:::|||   :::
Db        439 LeuCys------ThrGlyValLeuAsnIleSerIleValValProGlnMetAlaIleAla 456

Qy       1805 CTGTTTATGGCTCCATTGTCCAGCTC-----------------------------AGC 1834
                 |||      |||      :::|||
Db        457 LeuGlyAlaGlyProTrpAspGluLeuPheGlyGluGlyAsnIleProAlaPheAlaMet 476

Qy       1835 CAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTC 1882
                    ||||||   |||         :::|||||||||:::   |||:::
Db        477 AlaSerValPheAlaAlaAlaAlaAlaAlaGlyValValLeuLeu 492
```

```
RESULT 44
Q9M535
ID   Q9M535     PRELIMINARY;      PRT;   530 AA.
AC   Q9M535;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Sucrose transport protein.
OS   Euphorbia esula (Leafy spurge).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Malpighiales; Euphorbiaceae; Euphorbia.
OX   NCBI_TaxID=3993;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=UNDERGROUND ADVENTITIOUS BUDS;
RA   Anderson J.V., Horvath D.P.;
RT   "Identification of mRNAs expressed in underground adventitious buds of
RT   Euphorbia esula (leafy spurge).";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF242307; AAF65765.1; -.
DR   InterPro; IPR003662; sub_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   TIGRFAMs; TIGR01301; GPH_sucrose; 1.
SQ   SEQUENCE   530 AA;  55843 MW;  059C1ED3BB02D356 CRC64;

Alignment Scores:
Pred. No.:               4.94e-13      Length:         530
Score:                   327.50        Matches:        137
Percent Similarity:      40.64%        Conservative:   91
Best Local Similarity:   24.42%        Mismatches:     212
Query Match:             5.10%         Indels:         121
DB:                      10            Gaps:           19

US-09-759-143-110 (1-3410) x Q9M535 (1-530)
```

```
Db        772 laAspGlyGlnProGlyAlaLysGlyGluGlnGlyGluAlaGlyGlnLysGlyAspAlaG 792

Qy       1181 --GCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGCT 1124
                ||||||   |||:::   |||          |||            |||:::
Db        792 lyAlaProGlyProGlnGlyPro---SerGlyAlaProGlyProGlnGlyProThrGlyV 811

Qy       1123 GCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACAGCT---- 1068
               |||      |||:::   |||          |||      |||:::|||
Db        811 alThrGly---ProLysGlyAlaArgGlyAlaGlnGlyProProGlyAlaThrGlyPheP 830

Qy       1067 --GGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCC---GGAAAGCCAAGCGGGCCCGGCA 1013
                ||||||||||||      |||   |||||||||   |||   |||    ||||||
Db        830 roGlyAlaAlaGlyArgValGlyProProGlySerAsnGlyAsnProGlyProProGlyP 850

Qy       1012 TG---------------------------------------GACAGCAGTGGGGCGA 995
                                                        ||||||
Db        850 roProGlyProSerGlyLysAspGlyProLysGlyAlaArgGlyAspSerGlyProProG 870

Qy        994 CAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCTCCT---- 939
               |||   |||          |||          |||||||||   ||||||
Db        870 lyArgAlaGlyAspProGlyLeu-----------GlnGlyProAlaGlyProProGlyG 886

Qy        938 -----------------------------------------CAGCCACCAGCAGTGTGGC 920
                                                        ||||||
Db        886 luLysGlyGluProGlyAspAspGlyProSerGlyProAspGlyProPro---------- 902

Qy        919 TGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGCACTCCTCCTGGGTGCC 860
                              |||      |||   |||||||||||||||      |||    |
Db        903 -------------GlyProGlnGlyLeuAlaGlyGlnArgGlyIle---ValGlyLeuP 917

Qy        859 CAGGT-----------AGGGGG---CCAGGGCACTGGTGTCCCAGTCAATGGCAGGCAG 815
              |||||            ||||||   ||||||      ||||||
Db        917 roGlyGlnArgGlyGluArgGlyPheProGlyLeuProGlyProSer------------ 932

Qy        814 GAGGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGCG 755
                |||   ||||||                          :::|||      ||||||
Db        933 --GlyGluProGly--------------------------LysGlnGlyAlaProGlyA 943

Qy        754 ACAGTGGT-----------CCGGGT-----------CCCGGAAGAGGTCAGAGAGCAG 719
                ||||||             ||||||            ||||||         ||||
Db        943 laSerGlyAspArgGlyProProGlyProValGlyProProGlyLeuThrGlyProAlaG 963

Qy        718 GGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGCCCAGGA- 660
              ||   |||         |||||||||     |||                ||||||
Db        963 lyGluProGlyArgGluGlyThrProGlyAlaAspGlyPro-----------ProGlyA 979

Qy        659 -----------------------------------------TGAGCAGTGCCAGCTCCAG 641
                                                        |||:::   ||||||
Db        979 rgAspGlyAlaAlaGlyValLysGlyAspArgGlyGluAlaGlyAlaLeuGlyAlaProG 999

Qy        640 GGGCC-----------TGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGGCCCTTGG 593
              |||||                  ||||||   |||                |||    |
Db        999 lyAlaProGlyProProGlySerProGlyProAlaGly--------------ProThrG 1014

Qy        592 GA------TGAGAAAGAGGCTCAGCAGGATGCCCAAGGACAGTGCCCAGATGAAGGGCCG 539
              ||          :::||||||   |||||| ||||||||               ||||||
Db       1014 lyLysGlnGlyAspArgGlyGluAlaGly-AlaGlnGlyProMetGlyProAlaGlyPro 1033

Qy        538 GCGGC------------------------------------------GGCCATAGCGTC 522
               |||                                        |||      :::
Db       1034 AlaGlyAlaArgGlyLeuProGlyProGlnGlyProArgGlyAspLysGlyGluAlaGly 1053

Qy        521 CACGCCAGTGGTCACTGGCTGAGCCTAGGAGCGGGACACAGA------------------ 480
                   |||         |||        |||||||||
Db       1054 GluAlaGlyGluArgGlyLeuLys---------GlyHisArgGlyPheThrGlyLeuGln 1070

Qy        479 ---------------CCAGGCCCAGCACTGGACCAA---------TGCCCAGCA---CCA 447
                             |||||||||:::   ||||||        ||||||   |||
Db       1071 GlyLeuProGlyProProGlyProSerGlyAspGlnGlyAlaSerGlyProAlaGlyPro 1090

Qy        446 TGGTCATGAACTTCTCCTCTACCCCCACTT---CCAGCAGCAGAGGCGGCACATAGGTGA 390
                            |||   |||:::   |||:::   :::
Db       1091 SerGlyProArgGlyProProGlyProValGlyProSerGlyLysAspGlyAlaAsnGly 1110

Qy        389 TGCCTGCGGCCAAACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGG 330
                 |||       |||||||||   |||                      |||:::|||
Db       1111 IleProGlyProIleGlyProProGlyProArg--------------GlyArgSerGly 1125

Qy        329 CTTTCCGGT---GCCGCAGCAGGCGGCTCACCCACAGCCTCTGGACCATAGTGGGCCAGG 273
                 :::|||   |||          |||:::|||         ||||||   ||||||
Db       1126 GluThrGlyProAlaGlyProProGlyAsnProGlyProProGlyPro-ProGlyProPr 1145

Qy        272 CGGG-----------------------TAGGGCTCAGGGGGCCGTTCAGGCACTCCAGA 237
               |||                        |||   |||   |||         |||::
Db       1145 oGlyProGlyIleAspMetSerAlaPheAlaGlyLeuGlyProArgGluLysGlyProAs 1165

Qy        236 ACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAACT 177
               :   |||:::          |||   |||||||||
Db       1165 pProLeuGlnTyrMetArgAlaAspGluAlaAlaGly---------------------- 1177

Qy        176 GCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCA-------- 125
                                                      |||      |||
Db       1178 --------------------------------------GlyLeuArgProHisAspGl 1184

Qy        124 ---------GCTTCTCAGCCCATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTGGGG 75
                       |||:::         ||||||
Db       1184 uGluValGluAlaThrLeuLysSerLeuAsn-------------------------As 1195

Qy         74 ACACGTCTCATCACTCAGATCCTGGCCGAGGCGCGCGGCTGTCACCCGGAGCCAGCGCGT 15
                 :::   |||:::||||||         |||:::|||   :::|||   |||      |
Db       1195 nGlnIleGluSerIleArgSer-ProGluGlySerArgLysAsnPro---AlaArgThrC 1214

Qy         14 GCAGG 10
              |||||
Db       1214 ysArg 1215

RESULT 43
Q948L0
ID   Q948L0      PRELIMINARY;      PRT;   506 AA.
AC   Q948L0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Sucrose transporter.
GN   OSSUT3.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. NIPPONBARE; TISSUE=PANICLE;
RA   Hirose T., Scofield G.N., Whitfeld P.R., Aoki N., Furbank R.T.,
RA   Terao T.;
RT   "Isolation and characterisation of a cDNA for a novel sucrose
RT   transporter, OsSUT3, from rice.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB071809; BAB68368.1; -.
DR   InterPro; IPR003662; sub_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   TIGRFAMs; TIGR01301; GPH_sucrose; 1.
SQ   SEQUENCE   506 AA;  52776 MW;  0844DC10E1E63E75 CRC64;

Alignment Scores:
Pred. No.:               4.91e-13      Length:        506
Score:                   327.50        Matches:       143
Percent Similarity:      39.57%        Conservative:  77
Best Local Similarity:   25.72%        Mismatches:    207
Query Match:             5.10%         Indels:        129
DB:                      10            Gaps:          21

US-09-759-143-110 (1-3410) x Q948L0 (1-506)

Qy        335 CTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACC 394
                 |||   |||         ::::::      |||::::::         |||      ::::::
```

```
Qy   2912 AGAGAAGAGGGGTGGTTAGGGAAGCCGTTGAGACCTGAAGCCCCACCCTCTACCTTCCTT 2853
Db    162 GlyGluProGlyValSerGly---ProMetGlyProArgGlyProProGly-ProProGl 180
Qy   2852 CAACACCCTAACCTTGGGTAACAGCATTTGGAATTATCATTTGGGATGAGTAGAATTTCC 2793
Db    180 yLysPro--------------------------------------------------- 182
Qy   2792 AAGGTCCTGGGTTAGGCATTTTGGGGGGCCAGACCCCAGGAGAAGAAGATTCTGGCAATG 2733
Db    183 -------------------------GlyAspAspGlyGluAlaGlyLys----------- 190
Qy   2732 ATCAGCCCAATGACCAGCTATCTCAGGGGACCTGATTGTTGGGGATCCCCCACCCTACCC 2673
Db    191 ------------ProGlyLysSerGlyGluArg------GlyProProGlyPro----- 202
Qy   2672 AAATATTAGACACCAACACAGAAAAGCTAGCAATGGATTCCCTTCTACTTTGTTAAATAA 2613
Db    203 ---------------------GlnGlyAlaArgGlyPheProGlyThrProGlyLeuPr 215
Qy   2612 ATAAGTTAAATATTTAAATGCCTGTGTCTCTGTGATGGCAACAGAAGGACCAACAGGCCA 2553
Db    215 oGlyValLysGlyHisArgGlyTyrProGlyLeuAspGlyAlaLysGlyGluAlaGlyAl 235
Qy   2552 CATCCTGATAAAAGGTAAGAGGGGGGTGGATCAGCAAAAAGACAGTGCTGTGGGCTGAGG 2493
Db    235 aProGlyValLysGlyGluSerGlySerProGlyGluAsnGlySerProGlyProMetGl 255
Qy   2492 GGACCTGGTTCTTGTGTGTTGCCCCTCAGGACTCTTCCCCTACAAATAAGTCATATGTTC 2433
Db    255 y--------------------ProArgGlyLeu------------------------- 259
Qy   2432 AAATCCCATGGAGGAGTGTTTCATCCTAGAAACTCCCATGCAAGAGCTACATTAAACGAA 2373
Db    260 ----------------------------------ProGlyGluArgGlyArgThrGlyPr 268
Qy   2372 GCTGCAGGTTAAGGGGCTTAGA------GATGGGAAACCAGGTGACTGAGTTTATTCAGC 2319
Db    268 oAlaGlyAlaAlaGlyAlaArgGlyAsnAspGlyGlnProGly----------------- 282
Qy   2318 TCCCAAAAACCCTTCTCTAGGTGTGTCTCAACTAGGAGGCTAGCTGTTAACCCTGAGCCT 2259
Db    283 -ProAlaGlyProProGlyProValGlyProAlaGlyGly------------------Pr 296
Qy   2258 GGGTAATCCACCTGCAGAGTCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGT 2199
Db    296 oGlyPheProGlyAlaProGlyAlaLysGlyGluAla---GlyProThrGly-------- 312
Qy   2198 ATAAGTCCAGACTGAAACCCCCTTGGAAGGCCTCCAGTCA-------------------- 2159
Db    313 ---------AlaArgGlyPro---GluGlyAlaGlnGlyProArgGlyGluProGlyTh 328
Qy   2158 ----GGCAGCCCTAGAGACTGGGGAGAGAGG---------------------------- 2132
Db    328 rProGlySerProGlyProAlaGlyAlaAlaGlyAsnProGlyThrAspGlyIleProGl 348
Qy   2131 ------------------------------AGAGGGACGCCC------------------ 2120
Db    348 yAlaLysGlySerAlaGlyAlaProGlyIleAlaGlyAlaProGlyPheProGlyProAr 368
Qy   2119 -CAGCCCCCAGCTGTGCAGCTACGCACCTCA------GCAGCACAGGGTGGCAGCAGAGA 2067
Db    368 gGlyProProGlyProGlnGlyAlaThrGlyProLeuGlyProLysGlyGlnThrGlyGl 388
Qy   2066 GCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAA 2007
Db    388 uProGlyIleAlaGlyPheLysGlyGluGlnGlyProLysGlyGluProGlyProAlaGl 408
Qy   2006 CAGGAGCGGGGAGCTGGGACCCAGT---------------------------------- 1982
Db    408 yProGlnGlyAlaProGlyProAlaGlyGluGluGlyLysArgGlyAlaArgGlyGluPr 428
Qy   1981 -------------GAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTTCTACGCTGAGT 1935

Db    428 oGlyGlyAlaGlyProValGlyProProGlyGluArgGlyAlaProGlyAsnArgGlyPh 448
Qy   1934 ATTTGGCCAAGTCGCTCTTGTC------------------------------------- 1913
Db    448 eProGlyGlnAspGlyLeuAlaGlyProLysGlyAlaProGlyGluArgGlyProSerGl 468
Qy   1912 -------------AAATACTACCTGTGTAGCAAAGTAAATGGCGACCA------------ 1878
Db    468 yLeuAlaGlyProLys---------------GlyAlaAsnGlyAspProGlyArgProGl 483
Qy   1877 -GACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTGGACAATG 1819
Db    483 yGluProGlyLeu---------ProGlyAlaArgGlyLeuThrGlyArgProGlyAspAl 500
Qy   1818 GAGCCCATAAACAGGGATGGGGCCA---------CCTGGGACAGCAGGAAGG-------- 1776
Db    500 aGlyProGlnGlyLysValGlyProSerGlyAlaProGlyGluAspGlyArgProGlyPr 520
Qy   1775 ----------------------------------------CACTATCCAGGATGGCGAGG 1756
Db    520 oProGlyProGlnGlyAlaArgGlyGlnProGlyValMetGlyPheProGlyProLysGl 540
Qy   1755 TCCAGGCAGATGCCCCGGC--------------------CCGGAACCACCCTGGCCTCG 1717
Db    540 yAlaAsnGlyGluProGlyLysAlaGlyGluLysGlyLeuPro-GlyAlaProGlyLeuA 560
Qy   1716 GTGGGCTCACCCACCACCACACGTACGGAGACATCACAGGCAGAGGCCCCGCAGAGCGCG 1657
Db    560 rgGlyLeu-ProGlyLysAsp---GlyGluThrGlyAlaAlaGlyProProGlyProAla 578
Qy   1656 GGTGGAGGTGGGAGCAGG------------------------------------------ 1639
Db    579 GlyProAlaGlyGluArgGlyGluGlnGlyAlaProGlyProSerGlyPheGlnGlyLeu 598
Qy   1638 CCACTGCCTCCAGCACCCACGTGTCCATTAGGGAAG----------------------- 1603
Db    599 ProGlyProProGlyProProGlyGluGlyGlyLysProGlyAspGlnGlyValProGly 618
Qy   1602 ------GGAGCTCCAGGCTTA---GGGCCTGGCAGGAAGC-----------TGGTCATC 1564
Db    619 GluAlaGlyAlaProGlyLeuValGlyPro-ArgGlyGluArgGlyPheProGlyGluAr 638
Qy   1563 AGGCTGTCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGC 1504
Db    638 gGlySerProGlyAlaGlnGlyLeuGlnGlyAlaArgGlyLeu---ProGlyThrProGl 657
Qy   1503 TTCTCCCGGTGGTAGAGGGAGG---------------CCAGTGTGTAGGGCAGGATCTGC 1459
Db    657 yThrAspGlyProLysGlyAlaSerGlyProAlaGlyProProGlyAlaGlnGlyProPr 677
Qy   1458 AGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAGCTG-------TCACCACGGCCAC 1406
Db    677 oGlyLeu-GlnGlyMetProGlyGluArgGlyAlaAlaGlyIleAlaGlyProLysGlyA 697
Qy   1405 ACTGTGGGACAGGCATGTGGCACCGGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATA 1346
Db    697 spArgGly---AspValGlyGluLysGlyProGluGlyAlaProGlyLysAspGlyGlyA 716
Qy   1345 GACTGCTCGAGTGCCGAATCGCT------GCACCAGCCGGTCCATGACCAGAGAGA---- 1296
Db    716 rgGlyLeuThrGlyProIleGlyProProGlyProAlaGlyAlaAsnGlyGluLysGlyG 736
Qy   1295 --------AGACCAGGGAGATGGCGCACTGCAGGAACAGCCCCAGGCTGCCCATCCGAAC 1244
Db    736 luValGlyProProGlyProAlaGlyThrAlaGly---AlaArgGlyAlaProGlyGluA 755
Qy   1243 GCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTA 1184
Db    755 rg---------GlyGluThrGlyProProGlyProAlaGlyPheAlaGlyProProGlyA 772
Qy   1183 CA---------------------------------------------------------- 1182
```

```
Qy    398 --------------GTGCCGCCTCTGCTGCTGGAAGTGGGGTAGAGGAGAAGTTCATG 442
                        :::  |||  :::      :::|||:::          |||
Db     79 LeuGlnLeuSerLeuLeuThrProTyrIleGlnThrLeuGlyIleSerHisAlaPheSer 98

Qy    443 ACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCA 502
          :::   :::          ||||||:::   ||||||||      |||   :::|||
Db     99 SerPheIleTrpLeuCysGlyProIleThrGlyLeuValValGlnProPheValGlyIle 118

Qy    503 GCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCC 562
             ||||||            :::||||||||||||||||||||||        |||
Db    119 TrpSerAspLysCysThrSerLysTyrGlyArgArgArgProPheIleLeuValGlySer 138

Qy    563 TTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCA------GGG 616
             |||  :::::::::  :::|||          |||:::  |||     |||
Db    139 PheMetIleSerIleAlaValIleIleIle---------GlyPheSerAlaAspIleGly 155

Qy    617 CTGCTGTGCCCGGAT----------------------------CCCAGGCCCCTGGAG 646
             |||    |||                                 |||
Db    156 TyrLeuLeuGlyAspSerLysGluHisCysSerThrPheLysGlyThrArgThrArgAla 175

Qy    647 CTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACT 706
             :::   |||:::|||   |||||||||
Db    176 AlaValValPheIleIleGlyPheTrpLeuLeuAspLeuAlaAsnAsnThrValGlnGly 195

Qy    707 CCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTAC 766
          |||    |||||||||:::||||||          ||||||           |||
Db    196 ProAlaArgAlaLeuLeuAlaAspLeu---SerGlyProAspGlnArgAsnThrAlaAsn 214

Qy    767 TCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATT 826
          :::|||:::           :::::::::|||      ||||||:::      |||
Db    215 AlaValPheCysLeuTrpMetAlaIleGlyAsnIleLeuGlyPheSerAlaGlyAlaSer 234

Qy    827 -----------GACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 874
                     :::|||                |||:::|||  |||
Db    235 GlyLysTrpGlnGluTrp--------------PheProPheLeu---ThrSerArgAla 248

Qy    875 TGCCTCTTTGGC----------------------CTGCTCACCCTCATCTTCCTCACC 910
          |||                                 ||||||  :::::::::||||||||
Db    249 CysCysAlaAlaCysGlyAsnLeuLysAlaAlaPheLeuLeuAlaValValPheLeuThr 268

Qy    911 ---TGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCA 967
             |||      |||:::      |||:::|||                  :::|||
Db    269 IleCysThrLeuValThrIleTyrPheAlaLysGluIleProPheThrSerAsnLysPro 288

Qy    968 GCAGAA---GGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGC 1024
                         |||||||||   |||
Db    289 ThrArgIleGlnAspSerAlaProLeuLeuAspAsp------------------------ 300

Qy   1025 TTGGCTTTCCGGAACCTG------------------------------------------ 1042
          |||     :::  |||
Db    301 LeuGlnSerLysGlyLeuGluHisSerLysLeuAsnAsnGlyThrAlaAsnGlyIleLys 320

Qy   1042 ------------------------------------------------------------ 1042

Db    321 TyrGluArgValGluArgAspThrAspGluGlnPheGlyAsnSerGluAsnGluHisGln 340

Qy   1043 -----------------------GGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGC 1078
                                  |||:::::|||      |||
Db    341 AspGluThrTyrValAspGlyProGlySerValLeuValAsnLeuLeuThrSerLeuArg 360

Qy   1079 CGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTC 1138
             :::|||   :::      :::  :::         :::|||:::::
Db    361 HisLeuProProAlaMetHisSerValLeuIleValMetAlaLeuThrTrpLeuSerTrp 380

Qy   1139 ATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCC 1198
             |||  ||||||   ||||||:::::|||      :::|||  |||  |||
Db    381 PheProPhePheLeuPheAspThrAspTrpMetGlyArgGluValTyrHisGlyAspPro 400

Qy   1199 AGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGC 1258
             :::     :::        ||||||:::||||||||||   |||:::
Db    401 ThrGlyAsp------SerLeuHisMetGluLeuTyrAspGlnGlyValArgGluGlyAla 418

Qy   1259 CTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTG 1318
          |||||||||   |||        :::     :::   |||   :::::::::   :::
Db    419 LeuGlyLeuLeuLeuAsnSerValValLeuGlyIleSerSerPheLeuIleGluProMet 438

Qy   1319 GTGCAGCGATTCGGCACTCGAGCAGTCTAT---TTGGCCAGTGTGGCAGCTTTCCCTGTG 1375
             ||||||   |||  |||   |||:::   |||:::::           |||
Db    439 CysGlnArgMetGlyAlaArgValValTrpAlaLeuSerAsnPheThrValPheAlaCys 458

Qy   1376 GCTGCCGGTGCCACATGCCTGTCC------------------------------------ 1399
             ||||||      :::|||
Db    459 MetAlaGlyThrAlaValIleSerLeuMetSerLeuSerAspAspLysAsnGlyIleGlu 478

Qy   1400 --------------------------CACAGTGTGGCCGTGGTGACAGCTTCAGCCGCC 1432
                                    :::   ||||||:::           |||
Db    479 TyrIleMetArgGlyAsnGluThrThrArgThrAlaAlaValIleVal------PheAla 496

Qy   1433 CTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTAC 1492
             |||  ||||||     :::                   :::|||::::::::  :::
Db    497 LeuLeuGlyPheProLeuAlaIleThrTyrSerValProPheSerValThrAla------ 514

Qy   1493 CACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
                                       :::      |||:::||||||
Db    515 --------------------------GluValThrAlaAspSerGlyGly 522
```

RESULT 42
Q28396

```
ID   Q28396     PRELIMINARY;      PRT;  1418 AA.
AC   Q28396;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Type II collagen.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Richardson D.W., Dodge G.R.;
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 18-68 FROM N.A.
RA   MacLeod J.N., Fubini S.L., Gu D.N., Tetreault J.W., Todhunter R.J.;
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U62528; AAB05773.1; -.
DR   EMBL; AF040638; AAB96768.1; -.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR000885; Fib_collagen_C.
DR   Pfam; PF01410; COLFI; 1.
DR   Pfam; PF01391; Collagen; 18.
DR   ProDom; PD000007; Collagen; 4.
DR   ProDom; PD002078; Fib_collagen_C; 1.
DR   SMART; SM00038; COLFI; 1.
KW   Collagen.
SQ   SEQUENCE   1418 AA;  134343 MW;  115FCD19EB8696A3 CRC64;
```

Alignment Scores:

Pred. No.:	5.39e-13	Length:	1418
Score:	328.00	Matches:	338
Percent Similarity:	30.77%	Conservative:	58
Best Local Similarity:	26.26%	Mismatches:	413
Query Match:	5.27%	Indels:	481
DB:	6	Gaps:	73

US-09-759-143-110 (1-3410) x Q28396 (1-1418)

```
Qy   3032 GTGGGGAAAGTTGGGGGTAGGGGAAAGTTGGGGGTAGGGGAAATTTTGGGCAGTGCCTTC 2973
             :::|||   :::|||   ||||||      |||               |||
Db    134 MetGlyProMetGlyProArgGlyProProGlyProAlaGlyAlaProGlyPro------ 151

Qy   2972 ATCAGCCCAGTCCTAGAGAGAGTAGAGGGGAGTGGAAGTGGGGGGAACCAGGCTGGGCCA 2913
                                      |||   ||||||         |||
Db    152 ----------------------------GlnGlyPheGlnGlyAsnProGlyGluPro 161
```

```
Qy     575 CTGAGCCTCTTTCTCATC---------CCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGC 625
           :::|||:::   :::|||                |||:::    |||
Db     106 IleSerValIleValIleGlyPheSerSerAspIleGlyTyrAlaLeuGlyAspAlaThr 125

Qy     626 CCGGAT------------------CCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGC 667
              |||                   ||||||           |||     |||||||||
Db     126 GluAspCysLysValTyrArgGlyProArgTyrHisAlaAlaAlaAlaPheIleLeuGly 145

Qy     668 GTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCT 727
                 ||||||||||||                       |||     ||||||::::::
Db     146 PheTrpLeuLeuAspPheSerAsnAsnThrValGlnGlyProAlaArgAlaLeuMetAla 165

Qy     728 GACCTCTTC------CGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTC 781
           ||||||            |||                  |||    :::::::::
Db     166 AspLeuSerGlyArgHisGlyProSer---------AlaAlaAsnAlaIlePheCysSer 182

Qy     782 ATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATT---GACTGGGACACC 838
              ::::::||||||     |||||||||          :::      ||||||
Db     183 TrpMetAlaLeuGlyAsnIleLeuGlyTyrSerSerGlySerThrAsnAspTrp------ 200

Qy     839 AGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGC--------------------- 877
                          |||:::|||   |||:::      |||
Db     201 HisLysTrpPheProPheLeuMetThrArgAlaCysCysGluAlaCysAlaAsnLeuLys 220

Qy     878 CTCTTTGGCCTGCTCACCCTCATCTTCCTC---ACCTGCGTAGCAGCCACACTGCTGGTG 934
                    |||:::   ::::::||||||              |||   |||::::::
Db     221 AlaAlaPheLeuValAlaValValPheLeuGlyLeuSerThrAlaValThrMetValPhe 240

Qy     935 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 994
           |||   |||   ||||||   |||         |||:::
Db     241 AlaArgGluValAlaLeuAspProValAlaAlaAlaLys--------------------- 253

Qy     995 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGC--------- 1045
                                                   ||||||   |||
Db     254 ---------------------------------------ArgAsnGluGlyGluAlaSer 260

Qy    1046 ---------GCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTG 1096
                    |||:::         :::      |||      ||||||      |||
Db     261 GlyProLeuAlaValPheLysGlyMetLysAsnLeuProValGlyMetProSerValLeu 280

Qy    1097 CGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTT 1156
                       :::         :::|||::::::            |||   ||||||
Db     281 ------------IleValThrGlyLeuThrTrpLeuSerTrpPheProPheIleLeuPhe 296

Qy    1157 TACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACC 1216
              |||||||:::::|||         |||   |||   |||      |||
Db     297 AspThrAspTrpMetGlyArgGluIleTyrHisGlyArgProAspGlySerProAlaGlu 316

Qy    1217 GAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAG 1276
                    :::      ||||||||||||   |||:::   ||||||   |||
Db     317 ValThrAla------PheGlnGluGlyValArgGlnGlyAlaPheGlyLeuLeuLeuAsn 334

Qy    1277 TGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACT 1336
              :::      :::   |||   :::::::::   :::   :::|||   |||
Db     335 SerIleValLeuGlyIleSerSerPheLeuIleGluProMetCysArgArgLeuGlyAla 354

Qy    1337 CGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTG 1396
           |||||||||::::::   |||   |||         ||||||   |||
Db     355 ArgAlaValTrpValMetSerSerAlaValValCysValAlaMetAlaAla--------- 371

Qy    1397 TCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCC 1456
                    |||:::|||::::::|||
Db     372 ---------ValSerValLeuSerAla--------------------------------- 377

Qy    1457 CTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTG 1516
                          ::::::|||
Db     378 ---------------TrpSerLeu------------------------------------ 380

Qy    1517 CCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTC 1576
                       ||||||   ||||||   |||   :::|||:::
Db     381 ------------GlyAspPheGlyGly---SerValGlnAspAlaAlaArg--------- 392

Qy    1577 CTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGT 1636
                                 ||||||       |||   |||   |||
Db     393 ---------------------AlaProAlaGluGluGlyGlyValArgAlaSerAlaLeu 405

Qy    1637 GGCCTGCTCCCA------------CCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGAT 1684
              |||                   |||      ||||||
Db     406 AlaLeuPheValPheLeuGlyLeuProPheAlaValLeuCys------------------ 419

Qy    1685 GTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGC 1744
                ||||||      |||               |||         |||:::|||
Db     420 ---SerValProPheAlaValThrAlaGlnLeuThrAlaSerArgGlyGlyGlyGlnGly 438

Qy    1745 ATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCC 1804
           :::|||            :::|||:::   :::   ::::::   |||:::|||   :::
Db     439 LeuCys------ThrGlyValLeuAsnIleSerIleValValProGlnMetAlaIleAla 456

Qy    1805 CTGTTTATGGGCTCCATTGTCCAGCTC-----------------------------AGC 1834
              |||   |||         :::|||
Db     457 LeuGlyAlaGlyProTrpAspGluLeuPheGlyGluGlyAsnIleProAlaPheAlaMet 476

Qy    1835 CAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTC 1882
              ||||||   |||         :::||||||||||:::   |||:::
Db     477 AlaSerValPheAlaAlaAlaAlaAlaAlaAlaGlyValValLeuLeu 492
```

RESULT 41
O80605

```
ID   O80605      PRELIMINARY;      PRT;   594 AA.
AC   O80605;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Putative sucrose/H+ symporter (Sucrose transporter).
GN   T17M13.3 OR SUC3.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
RA   Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;
RT   "Arabidopsis thaliana chromosome II BAC T17M13 genomic sequence.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Meyer S., Truernit E., Huemmer C., Besenbeck R., Stadler R., Sauer N.;
RT   "AtSUC3: an unusual sucrose transporter from Arabidopsis.";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR   EMBL; AC004138; AAC32907.1; -.
DR   EMBL; AJ289165; CAB92307.1; -.
DR   InterPro; IPR003662; sub_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   TIGRFAMs; TIGR01301; GPH_sucrose; 1.
KW   Transmembrane.
SQ   SEQUENCE   594 AA;  63972 MW;  DF075A3601299DA0 CRC64;
```

Alignment Scores:

Pred. No.:	4.34e-13	Length:	594
Score:	328.50	Matches:	120
Percent Similarity:	38.63%	Conservative:	72
Best Local Similarity:	24.14%	Mismatches:	174
Query Match:	5.12%	Indels:	131
DB:	10	Gaps:	18

US-09-759-143-110 (1-3410) x O80605 (1-594)

```
Qy     347 AACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTAT--------- 397
           :::|||:::|||         |||            :::|||||||||:::   :::
Db      59 SerLeuValThrLeuValLeuSerCysThrValAlaAlaGlyValGlnPheGlyTrpAla 78
```

```
Db      640 nGlyAlaThrGly--GluProGlyLysAlaGlyGluGlnGlyGlyProGlyGluValGly 659

Qy     1151 ----------------------------------------GCGTGAAGGTCATGAGTG 1134
                                                    |||       :::
Db      660 GlyProGlyProSerGlySerArgGlyAspArgGlyPheProGlyGluArgGlyGlyIle 679

Qy     1133 CCATCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGTGCGGGGCATGCGGCAGC 1074
               |||:::       |||   ||||||   |||          |||     |||:::
Db      680 GlySerAlaGlyProAlaGlyProArgGlyAlaAsnGlySerProGlyAsnAspGlyAla 699

Qy     1073 ---------------ACAGCTGGTGCAGCCGGGGAAGCAGGGCGCCCAGGT---TCCGGA 1032
                           ||||||||   |||      |||   ||||||      |||
Db      700 ArgGlyGluSerGlyAlaAlaGlyAlaProGlyGlyMetGlyAlaProGlyLeuGlnGly 719

Qy     1031 AAGCCAAGC------GGGCCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACA 978
             |||         |||   |||   :::|||              ||||||
Db      720 MetProGlyGluArgGlySerGlyGlyAsnSerGlyAlaLysGlyGluArgGlyAspGly 739

Qy      977 GCCCTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCTCCTCAGCCACCAGCAGTGTGGCTG 918
                          |||     ||||||
Db      740 GlyProLysGlyAlaAspGlyGlyPro------------------------------ 748

Qy      917 CTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGCACTCCTCCTGGGTGCCCA 858
                          |||:::   |||     |||     |||   |||   |||   |||
Db      749 -----------GlyLysAspGlyMetArgGlyMetThrGly---ProIleGlyProPro 763

Qy      857 GGTAGGGGGCCAGGGCACTGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGC 798
             |||          |||           :::         ||||||   ||||||
Db      764 GlyProThrGlyAlaHisGlyGluLysGlyGluGlyGlyLeuGlyGlyProProGlyPro 783

Qy      797 CCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGTCCG------ 744
                                    :::     ||||||:::   |||
Db      784 ThrGlyGly-----------------ArgGlySerProGlyGluArgGlyGluHisGly 797

Qy      743 GGTCCCGGAAGAGGTCAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGA 684
                ||||||            |||||||||||              ||||||::::::
Db      798 AlaProGlyProAlaGlyPheAlaGlyProProGlyAlaAspGlyGlnProGlyAsnLys 817

Qy      683 ------------------------------AGTCCAGCAGCCCCACGCCCAGGATGAGCA 654
                                                ||||||    ||||||
Db      818 GlyGluThrGlyAsnAsnGlyProLysGlyGluAlaGlyAlaProGlyProGlyGlyPro 837

Qy      653 GTGCCAGCTCCAGGGGCCTGGGATCCG---GGCACAGCAGCCCTGCTAGCCAGCCGGCCC 597
             |||   |||||||||         |||   |||       :::            :::|||
Db      838 ValGlyAlaProGlyProGlnGlyProAlaGlyAsnSerGlyThrLysGlyThrArgGly 857

Qy      596 TTGGGATGAGAAAGAGGCTCAGCAGGATGCCCAAGGACAGTGCCCAGATGAAGGGCCGGC 537
                           |||:::::::|||    |||                       |||
Db      858 AlaProGlyProProGlyAlaSerGlyMetPro----------------------Gly 869

Qy      536 GGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTAGGAGCGGGACACAGACCA 477
                |||    |||           |||               |||:::|||      |||
Db      870 ProGlyGlyArgValGlyProProGly---------------GlySerGly---AlaPro 883

Qy      476 GGCCCAGCACTGGACCAATGCCCAGCACCATGGTCATGAACTTCTCCTCTACCCCCACTT 417
             |||                                        |||    ||||||
Db      884 Gly---------------------------------------SerAlaGlyProProGly 890

Qy      416 CCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCCAAACACACCTCCAGGCCAAAGG 357
             ||||||   :::        :::                              |||   |||
Db      891 ProAlaGlyLysGluGlyGlnLys--------------------------GlyGlyArg 901

Qy      356 -----------TTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGT----------- 321
                        :::||||||   |||      ||||||   :::|||
Db      902 GlyGluThrGlyIleAlaGlyArgProGlyGluAlaGlyAlaAlaGlyProProGlyPro 921

Qy      320 ---GCCGCAGCAGGC-----------------------GGCTCACCCACAGCCTCTGGA 288
                       |||:::|||                         |||:::|||      |||
Db      922 SerGlyAlaSerGlyAlaLysGlyAsnAspGlyProMetGlyAlaProGlyThrProGly 941

Qy      287 CCATAGT-----------------------------GGGCCAGGCGGGTAGGGCTCAGG 258
             |||                                   ||||||||||||   |||
Db      942 Pro-GlyGlyIleAlaGlyGlnArgGlyIleValGlyGlyProGlyGlyArgGlyProSe 961

Qy      257 GGGCCGTTCAGGCACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTC 198
                |||    :::|||   |||                |||:::|||    |||
Db      961 rGlyThrAlaGlyIlePro---------------GlyThrAlaGlyGluProGlyLysGl 976

Qy      197 TCCTCCTTGCTGCCGCCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTT 138
                   |||                  |||      ||||||                  |||
Db      976 nGlyProGlyGlyProValGlyGluArgGlySerProGly--------------ProMe 991

Qy      137 TGGTGCCGGTCCAGCTTCTCAGCCCATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTG 78
               ||| |||       |||                     |||        |||         |
Db      991 tGly-ProProGlyLeu---------SerGlyAlaProGlyGluAlaGlyArg---GluG 1007

Qy       77 GGGACACGTCTCATCACTCAGATCCTGGCCGA-------------------------- 46
             ||   |||   |||               ||||||
Db     1007 lySerThrGlyHisAspGlyValSerGlyArgAspGlyProProGlyProLysGlyAspA 1027

Qy       45 -----------GGCGCGCGGCTGTCACCCGGAGCC 22
                        |||         ||||||||||||
Db     1027 rgGlyGluAsnGlyAsnAlaGlySerProGlyAla 1038
```

RESULT 40
Q944W2

```
ID   Q944W2      PRELIMINARY;      PRT;   506 AA.
AC   Q944W2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Sucrose transporter.
OS   Oryza sativa (indica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39946;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. IR 36;
RA   Whitfeld P.R., Aoki N., Scofield G.N., Hirose T., Terao T.,
RA   Furbank R.T.;
RT   "Isolation and characterisation of a putative sucrose transporter gene
RT   (OsSUT3) from rice.";
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF419298; AAL14982.1; -.
DR   InterPro; IPR003662; sub_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   TIGRFAMs; TIGR01301; GPH_sucrose; 1.
SQ   SEQUENCE   506 AA;  52774 MW;  A4E73029D7022B64 CRC64;
```

Alignment Scores:

Pred. No.:	4.23e-13	Length:	506
Score:	328.50	Matches:	143
Percent Similarity:	39.57%	Conservative:	77
Best Local Similarity:	25.72%	Mismatches:	207
Query Match:	5.12%	Indels:	129
DB:	10	Gaps:	21

US-09-759-143-110 (1-3410) x Q944W2 (1-506)

```
Qy      335 CTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACC 394
               |||   |||        ::::::      |||:::::::          |||      ::::::
Db       26 LeuPheLeuAlaCysMetValAlaGlyGlyValGlnTyrGlyTrpAlaLeuGlnLeuSer 45

Qy      395 TATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTG 454
                :::   |||   :::      :::|||:::                   :::::::::
Db       46 LeuLeuThrProTyrIleGlnThrLeuGlyIleProHisAlaLeuThrSerValMetTrp 65

Qy      455 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 514
                      ||||||:::   ||||||:::      |||    :::|||      ||||||
Db       66 LeuCysGlyProIleAlaGlyLeuIleValGlnProCysValGlyLeuTyrSerAspLys 85

Qy      515 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 574
                          |||||||||||||||||||||                   :::   |||
Db       86 CysThrSerSerLeuGlyArgArgArgProPheIleLeuThrGlyCysIleIleIleCys 105
```

```
Db        474 lAlaAlaThrValSerGlyIleIleAlaLeuThrAlaLeuProSerProProSer----- 492

Qy       1666 CGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGGT 1708
                                  ||||||   :::           |||
Db        493 ----------------AspValLysIleLeuAlaThrGlyGly 501

RESULT 39
Q910B9
ID   Q910B9      PRELIMINARY;      PRT;  1458 AA.
AC   Q910B9;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Collagen a3(I).
GN   COL1A3.
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8022;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21257802; PubMed=11358497;
RA   Saito M., Takenouchi Y., Kunisaki N., Kimura S.;
RT   "Complete primary structure of rainbow trout type I collagen
RT   consisting of a1(I)a2(I)a3(I) heterotrimers.";
RL   Eur. J. Biochem. 268:2817-2827(2001).
DR   EMBL; AB052836; BAB55662.1; -.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR000885; Fib_collagen_C.
DR   InterPro; IPR001007; VWF_C.
DR   Pfam; PF01410; COLFI; 1.
DR   Pfam; PF01391; Collagen; 18.
DR   ProDom; PD000007; Collagen; 1.
DR   ProDom; PD002078; Fib_collagen_C; 1.
DR   PROSITE; PS01208; VWFC; UNKNOWN_1.
SQ   SEQUENCE   1458 AA;  137758 MW;  AB1F9F3410A98650 CRC64;

Alignment Scores:
Pred. No.:              3.74e-13      Length:         1458
Score:                  330.50        Matches:        247
Percent Similarity:     30.72%        Conservative:   47
Best Local Similarity:  25.81%        Mismatches:     320
Query Match:            5.31%         Indels:         344
DB:                     13            Gaps:           47

US-09-759-143-110 (1-3410) x Q910B9 (1-1458)

Qy       2365 GTTAAGGGGCTTAGA---------------GATGGGAAACCAGGTGACTGAGTTTATTCA 2321
               :::||||||   |||                ||||||     ||||||         :::
Db        255 IleLysGlyHisArgGlyPheAsnGlyMetAspGlyAlaLysGlyAspGlyGlyProAla 274

Qy       2320 GCTCCCAAAAACCCTTCTCTAGGTGTGTCTCAACTAGGAGGCTAGCTGTTAACCCTGAGC 2261
                    ||||||
Db        275 GlyProLysGlyGlu--------------------------------------------- 279

Qy       2260 CTGGGTAATCCACCTGCAGAGTCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCT 2201
                 |||      |||      |||   :::|||   |||              |||||||||
Db        280 ---GlyGlySerProGlySerSerGlyAsnSerGlyAlaMetGlyAlaArgGlyLeuPro 298

Qy       2200 GTATAAGTCCAGACTGAAACCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGG 2141

Db        298 ------------------------------------------------------------ 298

Qy       2140 GGAGAGAGGAGAGGGACGCCCCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAG 2081
              ||||||   ||||||   |||   |||                   :::|||  |||:::
Db        299 GlyGlu---ArgGlyArgProGlyProAlaGly------------SerSerGlyAlaArg 313

Qy       2080 GGTGGCAGCAGAGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAG 2021
              |||                       |||||||||            |||:::
Db        314 GlyAsnAspGlyAsnSerGlyProAlaGlySerAsnGlyProThrGlyProSerGlyAla 333

Qy       2020 ---------CCCCATGGGGCTAACAGGAGCGGGGAGCTGGGACCC------------AGT 1982
                                 |||      ||||||           ||||||   ||||||
Db        334 ProGlyPheProGlyGlyAlaGlyAlaLysGlyGluThrGlyProGlnGlyGlyArgGly 353

Qy       1981 GAGGCAGGCCCT---------------------------------CCACCCCAATGTGCT 1955
                    ||||||                                      |||     |||
Db        354 SerAspGlyProGlnGlySerArgGlyGluProGlyAsnProGlyProAlaGlyAlaAla 373

Qy       1954 GGAAGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGC 1895
              |||                                                      |||
Db        374 GlyPro---------------------------------------------------Ser 376

Qy       1894 AAAGTAAATGGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGA------- 1842
                       |||         ||||||   |||:::        |||:::
Db        377 GlyAlaProGlyGlyAspGlyGlnAlaGlyGlyLysGlySerThrGlyAlaAlaGlyIle 396

Qy       1841 ------------------------CAGACTGGCTGAGCTGGACAATGGAGCCCATAAACA 1806
                                         |||   ||||||
Db        397 SerGlyAlaProGlyPheProGlyThrArgGlyProAlaGlyProProGlySerGlyGly 416

Qy       1805 GGGATGGGGCCACCTGGGACAGCAGGAAGGCAC--------------------------- 1773
              |||   ||||||   |||      |||   |||
Db        417 GlyAsnGlyProLysGlyAsnHisGlyGluHisGlyAlaGlnGlyProLysGlyGluPro 436

Qy       1772 ------------TATCCAGGATGGCGAGGTCCAGGCAGATGCCCCGGCCCG--------- 1734
                          ||||||      |||||||||      |||||||||
Db        437 GlyValLysGlyGluProGlySerAlaGlyProGlyGlyProProGlyProSerGlyGlu 456

Qy       1733 ------------------------GAACCACCCTGGCCTCGGTGGGCTCACCCACCACCA 1698
                                       ||||||   |||   ||||||:::
Db        457 GluGlyLysArgGlyAlaArgGlyGluPro-GlyGlyAlaGlyGlyVal----------- 472

Qy       1697 CACGTACGGAGACATCACAGGCAGAGGCCCCGCAGAGCGCGGGTGGAGGTGGGAGCAGGC 1638
                                          |||   |||||||||      |||   |||
Db        473 ---------------------GlyProProGlyGluArgGlyMet-ProGlyMetArgG 485

Qy       1637 CACTGCCTCCAGCACCCACGTGTCCATTAGGGAAGGGAGCTCCAGGC------------- 1591
                   |||   :::                |||||||||   ||||||
Db        485 lyPheProGlySerAspGlyAlaAlaGlyGlyLysGlyGlyProGlyGluArgGlyAlaA 505

Qy       1590 --------TTAGGG----------------------------CCTGGCAGGAAGCTGG 1569
                      :::|||                              ||||||         :
Db        505 snGlyProMetGlyAlaGlnGlyAlaThrGlyGluAlaGlyArgProGlyGluProGlyM 525

Qy       1568 TCATCAGGCTGTCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGTATT------------ 1521
              ::                :::           |||  ||||||  |||
Db        525 etProGlySerLysGlyMetThrGlySerProGlySerPro-GlyProAspGlyLysVal 544

Qy       1520 ----------------------------TGGGCAGGAACACCTGCTTCTCCCGGTGGT 1491
                                          :::|||   |||   |||   |||
Db        545 GlyProAlaGlyValThrGlyAsnAspGlyArgSerGlyProProGlySerSerGlyAla 564

Qy       1490 AGAGGGAGGCCAGTGTGTAGGGCA---------------GGATCTGCAGGGCTGAGAAGG 1446
              ||||||:::|||                              |||      |||   ||
Db        565 ArgGlyGlnProGlyValMetGlyPheProGlyProLysGlyProGly-GlyAspAsnGl 584

Qy       1445 TGAACCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGG 1386
              |:::|||||||||   |||                            |||      ||
Db        584 yLysProGlyGluArgGly--------------------GluAlaGlyAlaValGlyGl 597

Qy       1385 CACCGGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATC 1326
              |   |||:::                  |||:::          |||   |||
Db        597 yLeuGlyAla------------------ProGlyLysAspGlyAspSerGlyAlaProGl 611

Qy       1325 GCTGCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACA 1266
                    |||         |||   |||   :::|||:::|||
Db        611 yProGlnGlyProAlaGlySerGlnGlyGlu---LysGlyGluGlyGly----------- 626

Qy       1265 GCCCCAGGCTGCCCATCCGAACGCCTTCATCATAGTGTCTCCGGGCCTC---GGTGCCCG 1209
                    |||   ||||||   |||                ||||||   |||    ::
Db        627 -------ProAlaGlyProAsnGlyPhe------------GlnGlyLeuProGlyThrGl 640

Qy       1208 GCTCAGCTCTGGGCACGCCCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACA--- 1152
              :   ::::::|||    ||||||   |||                |||    :::
```

```
Qy        286 CATAGTGGGCCAGGCGGG----------------------TAGGGCTCAGGGGGCCGT 251
              ||    ||||||   |||                      |||   |||   |||
Db       1210 ro-ProGlyProProGlyProGlyIleAspMetSerAlaPheAlaGlyLeuGlyGlnArg 1229

Qy        250 TCAGGCACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCG 203
                         |||:::   |||:::        |||    :::||||||
Db       1230 GluLysGlyProAspProLeuGlnTyrMetArgAlaAspGlnAlaAla 1245

RESULT 38
Q9SP14
ID   Q9SP14      PRELIMINARY;      PRT;   502 AA.
AC   Q9SP14;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Sucrose transporter 1.
GN   SUT1.
OS   Alonsoa meridionalis.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Lamiales; Scrophulariaceae; Alonsoa.
OX   NCBI_TaxID=106274;
RN   [1]
RP   SEQUENCE FROM N.A..
RX   MEDLINE=21414505; PubMed=11523659;
RA   Knop C., Voitsekhovskaja O., Lohaus G.;
RT   "Sucrose transporters in two members of the Scrophulariaceae with
RT   different types of transport sugar.";
RL   Planta 213:80-91(2001).
DR   EMBL; AF191025; AAF04295.2; -.
DR   InterPro; IPR003662; sub_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
SQ   SEQUENCE   502 AA;  52327 MW;  E754CD030536522D CRC64;

Alignment Scores:
Pred. No.:              2.51e-13      Length:         502
Score:                  332.00        Matches:        137
Percent Similarity:     40.00%        Conservative:   77
Best Local Similarity:  25.61%        Mismatches:     208
Query Match:            5.17%         Indels:         113
DB:                     10            Gaps:           20

US-09-759-143-110 (1-3410) x Q9SP14 (1-502)

Qy        307 CCGCCTGCT---GCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGG 363
              |||||||||   |||||||||              :::::::::|||       :::     ||
Db         13 ProProAlaGlnAlaAlaProValLys-AsnIleIleValValAlaAlaIleAlaAlaGl 32

Qy        364 CCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGG 423
              |::::::         |||          :::    :::   |||    :::        :::||
Db         32 yValGlnPheGlyTrpAlaLeuGlnLeuSerLeuLeuThrProTyrValGlnLeuLeuGl 52

Qy        424 GGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTG 483
               ||||          :::            |||||||:::    |||||||:::
Db         52 yValProHisValTrpAlaAlaPheIleTrpLeuCysGlyProIleSerGlyLeuLeuVa 72

Qy        484 TGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC 543
                 |||::::::|||      ||||||:::           |||:::||||||||||||||
Db         72 lGlnProIleValGlyTyrTyrSerAspAsnCysThrSerArgPheGlyArgArgArgPr 92

Qy        544 CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGG 603
              ||||||   |||     :::         :::     :::::::::||||||||         ||
Db         92 oPheIleAlaAlaGlyAlaGlyLeuValAlaValAlaValPheLeuIle---------Gl 109

Qy        604 CTGGCTAGCAGGGCTG--------------------CTGTGCCCGGATCCCAGGCCCCT 642
              |:::    |||      |||                     |||              :::|||
Db        109 yPheAlaAlaAspLeuGlyHisMetGlyGlyAspSerLeuGlyLysGlyThrLysProAr 129

Qy        643 GGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTT 702
                 :::|||:::   ::::::|||          :::||||||              :::
Db        129 gAlaValAlaValPheValValGlyPheTrpIleLeuAspValAlaAsnAsnMetLeuGl 149

Qy        703 CACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGG---GACCCGGACCACTGTCGCCA 759
```

```
                 |||          |||||||||:::|||:::                  :::
Db        149 nGlyProCysArgAlaLeuLeuAlaAspMetSerAlaGlyAsnAlaLysLysMetSerSe 169

Qy        760 GGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTAC-------- 811
                |||    |||::::::::|||    :::::::::|||     ||||||||
Db        169 rAlaAsnSerMetPheSerPhePheMetAlaValGlyAsnValLeuGlyTyrAlaAlaGl 189

Qy        812 ------------CTCCTGCCTGCCATTGACTGGGAC---ACCAGTGCCCTGGCCCCCTA 855
                          |||          :::    |||    |||    |||    |||   ||
Db        189 ySerTyrThrArgLeuTyrLysValPheProPheSerLysThrGluAlaCysAlaValTy 209

Qy        856 CCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGT 915
              |         :::    |||    |||        :::    |||:::             ::
Db        209 rCysAlaAsnLeuLysSerCysPhePheIleSerValAlaLeuLeuLeuThrValThrIl 229

Qy        916 AGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGG 975
              :    |||      :::|||    |||                    |||    |||
Db        229 eIleAlaLeuCysIleValArgGluThrProTyrThrAlaProProGlu----------- 245

Qy        976 GCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCG 1035
                   |||    ::::::          |||          |||          |||
Db        246 ----GluAlaGlyThrValLysLysHisThrValProValPheGlyGluLeu-------- 261

Qy       1036 GAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCT 1095
                   |||||||||                                :::||||||   ::
Db        262 ----PheGlyAlaLeuLysAsp----------------------LeuProArgProMe 272

Qy       1096 GCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTT 1155
              :       |||    :::                :::|||:::|||          |||   |||||
Db        272 tTrpMetLeuLeuLeuValThrAlaLeuAsnTrpValAlaTrpPheProPheLeuLeuPh 292

Qy       1156 TTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCAC 1215
              |    ||||||:::::|||:::    :::|||    |||          :::          |||::
Db        292 eAspThrAspTrpMetGlyLysGluValTyrGlyGly------LysValGlyGluGlySe 310

Qy       1216 CGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCA 1275
              :                 |||||:::||||||||:::|||:::||||||||    |||
Db        310 r------------LeuTyrAspLysGlyValArgValGlyAlaLeuGlyLeuMetLeuAs 326

Qy       1276 GTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGC-- 1333
                 :::    |||    ||||||          :::          |||
Db        326 nSerValValLeuGlyValAlaSerLeuGlyValGlnValThrAlaArgGlyLeuGlyGl 346

Qy       1334 ------------------------------ACTCGAGCAGTCTATTTGGCCAGTGTGGC 1362
                                                 |||:::    ||||||
Db        346 yValLysArgLeuTrpGlySerValAsnLeuLeuLeuAlaIleCysLeuAlaMetThrVa 366

Qy       1363 AGCTTTC--------------------------------CCTGTGGCTGCCGGTGCCAC 1389
                                                      |||          |||
Db        366 lLeuIleThrLysMetAlaGlnHisHisArgGluTyrAlaSerValGlyGlyAlaAlaPr 386

Qy       1390 ATGCCTGTCC---------CACAGTGTGGCCGTG---GTGACAGCTTCAGCCGCCCTCAC 1437
                   |||          |||    |||    :::          |||    ::::::
Db        386 oThrLeuGlnProProValHisGlyValLysIleGlyAlaLeuAlaLeuPheSerValLe 406

Qy       1438 CGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCG 1497
                |||          :::|||          :::|||:::    |||||||||:::|||
Db        406 uGlyIleProLeuAlaAlaThrPheSerIleProPheAlaLeuAlaSerIleTyr----- 424

Qy       1498 GGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGA 1557
                                                         ||||||
Db        425 -----------------------------SerSerAsnSerGlyAlaGlyGlnGlyLe 434

Qy       1558 CAGCCTG------------------------------------ATGACCAGCTTCCTGCC 1581
               ||||||                                          |||
Db        434 uSerLeuGlyValLeuAsnLeuAlaIleValIleProGlnMetPheValSerValAlaSe 454

Qy       1582 AGGCCCTAAGCCTGGA--------------------GCTCCCTTCCCTAATGGACACGT 1620
               ||||||                                  |||    |||    |||
Db        454 rGlyProTrpAspAlaLeuPheGlyGlyGlyAsnLeuProAlaPheValValGlyAlaVa 474

Qy       1621 GGGTGCTGGAGGCAGTGGCCTG--------------CTCCCACCTCCACCCGCGCTCTG 1665
              |    |||    ||||||:::                  ||||||    ||||||:::
```

```
Db    531 yGluArgGlyProSerGlyLeuAlaGlyProLys--------------GlyAlaAsnGl 546
Qy   1883 CGACCA------------GACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGAC 1837
Db    546 yAspProGlyArgProGlyGluProGlyLeu---------ProGlyAlaArgGlyLeuTh 563
Qy   1836 TGGCTGAGCTGGACAATGGAGCCCATAAACAGGGATGGGGCCA---------CCTGGGAC 1786
Db    563 rGlyArgProGlyAspAlaGlyProGlnGlyLysValGlyProSerGlyAlaProGlyGl 583
Qy   1785 AGCAGGAAGG------------------------------------------CA 1774
Db    583 uAspGlyArgProGlyProProGlyProGlnGlyAlaArgGlyGlnProGlyValMetGl 603
Qy   1773 CTATCCAGGATGGCGAGGTCCAGGCAGATGCCCCGGC--------------------CC 1735
Db    603 yPheProGlyProLysGlyAlaAsnGlyGluProGlyLysAlaGlyGluLysGlyLeuPr 623
Qy   1734 GGAACCACCCTGGCCTCGGTGGGCTCACCCACCACCACACGTACGGAGACATCACAGGCA 1675
Db    623 o-GlyAlaProGlyLeuArgGlyLeu-ProGlyLysAsp---GlyGluThrGlyAlaAla 641
Qy   1674 GAGGCCCCGCAGAGCGCGGGTGGAGGTGGGAGCAGG----------------------- 1639
Db    642 GlyProProGlyProAlaGlyProAlaGlyGluArgGlyGluGlnGlyAlaProGlyPro 661
Qy   1638 -----------------CCACTGCCTCCAGCACCCACGTGTCCATTAGGGAAG------ 1603
Db    662 SerGlyPheGlnGlyLeuProGlyProProGlyProProGlyGluGlyGlyLysProGly 681
Qy   1602 -----------------------GGAGCTCCAGGCTTA---GGGCCTGGCAGGAAGC-- 1572
Db    682 AspGlnGlyValProGlyGluAlaGlyAlaProGlyLeuValGlyPro-ArgGlyGluAr 701
Qy   1571 ---------TGGTCATCAGGCTGTCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGTAT 1522
Db    701 gGlyPheProGlyGluArgGlySerProGlyAlaGlnGlyLeuGlnGlyProArgGlyLe 721
Qy   1521 TTGGGCAGGAACACCTGCTTCTCCCGCTGGTAGAGGGAGG--------------CCAGT 1477
Db    721 u---ProGlyThrProGlyThrAspGlyProLysGlyAlaSerGlyProAlaGlyProPr 740
Qy   1476 GTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAGCTG-- 1419
Db    740 oGlyAlaGlnGlyProProGlyLeu-GlnGlyMetProGlyGluArgGlyAlaAlaGlyI 760
Qy   1418 -----TCACCACGGCCACACTGTGGGACAGGCATGTGGCACCGGCAGCCACAGGGAAAGC 1364
Db    760 leAlaGlyProLysGlyAspArgGly---AspValGlyGluLysGlyProGluGlyAlaP 779
Qy   1363 TGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCT------GCACCAGCCGGTC 1310
Db    779 roGlyLysAspGlyGlyArgGlyLeuThrGlyProIleGlyProProGlyProAlaGlyA 799
Qy   1309 CATGACCAGAGAGA-----------AGACCAGGGAGATGGCGCACTGCAGGAACAGCCC 1262
Db    799 laAsnGlyGluLysGlyGluValGlyProProGlyProAlaGlyThrAlaGly---AlaA 818
Qy   1261 CAGGCTGCCCATCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGC 1202
Db    818 rgGlyAlaProGlyGluArg---------GlyGluThrGlyProProGlyProAlaGlyP 835
Qy   1201 TCTGGGCACGCCCTGGTACA------------------------------------- 1182
Db    835 heAlaGlyProProGlyAlaAspGlyGlnProGlyAlaLysGlyGluGlnGlyGluAlaG 855
Qy   1181 -------------------GCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGT 1142
Db    855 lyGlnLysGlyAspAlaGlyAlaProGlyProGlnGlyPro---SerGlyAlaProGlyP 874
Qy   1141 CATGAGTGCCATCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGTGCGGGGCAT 1082
Db    874 roGlnGlyProThrGlyValThrGly---ProLysGlyAlaArgGlyAlaGlnGlyProP 893
```

```
Qy   1081 GCGGCAGCACAGCT------GGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCC---GGAA 1031
Db    893 roGlyAlaThrGlyPheProGlyAlaAlaGlyArgValGlyProProGlySerAsnGlyA 913
Qy   1030 AGCCAAGCGGGCCCGGCATG----------------------------------------- 1011
Db    913 snProGlyProProGlyProProGlyProSerGlyLysAspGlyProLysGlyAlaArgG 933
Qy   1010 --GACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCC 953
Db    933 lyAspSerGlyProProGlyArgAlaGlyAspProGlyLeu-----------GlnGlyP 949
Qy    952 CAGCGCTGCCTCCT--------------------------------------------C 938
Db    949 roAlaGlyProProGlyGluLysGlyGluProGlyAspAspGlyProSerGlyProAspG 969
Qy    937 AGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAG 878
Db    969 lyProPro-----------------------GlyProGlnGlyLeuAlaGlyGlnArgG 981
Qy    877 GCACTCCT-------------------------------CCTGGGTGCCCAGGT---- 855
Db    981 lyIleValGlyLeuProGlyGlnArgGlyGluArgGlyPheProGlyLeuProGlyProS 1001
Qy    854 ----------------------------------------AGGGGG---------- 849
Db   1001 erGlyGluProGlyLysGlnGlyAlaProGlyAlaSerGlyAspArgGlyProProGlyP 1021
Qy    848 ----------CCAGGGCACTGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCA 800
Db   1021 roValGlyProProGlyLeuThrGlyProSer--------------GlyGluProGly- 1035
Qy    799 GCCCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGTCCGGGTC 740
Db   1036 ------------------------ArgGluGlySerProGlyAlaAspGlyPro---P 1046
Qy    739 CCGGAAGAGGTCAGAGAGCAGGG------------------------CCTCCAGTGGAGT 704
Db   1046 roGlyArgAspGlyAlaAlaGlyValLysGlyAspArgGlyGluThrGlyProVal---- 1064
Qy    703 GAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCA---CGCCCAGGATGAGCAGTGCCAG 647
Db   1065 --GlyAlaProGly------AlaProGlySerProGlySerProGlyProAlaGlyProT 1082
Qy    646 CTCCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAG 587
Db   1082 hrGlyLysGlnGlyAspArgGlyGluAlaGly------AlaGlnGlyProMetGly---P 1099
Qy    586 AAAGAGGCTCAGCAGGATGCCCAAGGACAGTGCCC---------AGATGAAGGGCCGGCG 536
Db   1099 roAlaGlyProAlaGlyAla---ArgGlyIleProGlyProGlnGlyProArgGlyAspL 1118
Qy    535 GCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTAGGAGCGGGACACAGA---- 480
Db   1118 ysGlyGluAlaGlyGluAlaGlyGluArgGlyLeuLys---------GlyHisArgGlyP 1135
Qy    479 ----------------------------CCAGGCCCAGCACTGGACCAA---------T 458
Db   1135 heThrGlyLeuGlnGlyLeuProGlyProProGlyProSerGlyAspGlnGlyAlaSerG 1155
Qy    457 GCCCAGCA---CCATGGTCATGAACTTCTCCTCTACCCCCACTT---CCAGCAGCAGAGG 404
Db   1155 lyProAlaGlyProSerGlyProArgGlyProProGlyProValGlyProSerGlyLysA 1175
Qy    403 CGGCACATAGGTGATGCCTGCGGCCAAACACACCTCCAGGCCAAAGGTTAGCAGGTTGAC 344
Db   1175 spGlyAlaAsnGlyIleProGlyProIleGlyProProGlyProArg------------ 1190
Qy    343 CAGCAAGAGCTGGGCTTTCCGGT---GCCGCAGCAGGCGGCTCACCCACAGCCTCTGGAC 287
Db   1191 --GlyArgSerGlyGluThrGlyProAlaGlyProProGlyAsnProGlyProProGlyP 1210
```

```
Db        454 SerValGlyGlyGlyProPheAspGluLeuPheHisGlyGlyAsnIleProAla------ 471

Qy       1592 CCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTC------ 1645
                           |||     |||   :::   |||     |||||||:::|||
Db        472 -----------PheValLeuGlyAlaIleAlaAlaAlaValSerGlyIleLeuAlaLeu 487

Qy       1646 ---------------CCACCTCCACCCGCG 1660
                             |||||||||   |||
Db        488 ThrValLeuProSerProProProAspAla 497

RESULT 37
O77753
ID   O77753      PRELIMINARY;      PRT;  1487 AA.
AC   O77753;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Type IIA procollagen.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98340920; PubMed=9676231;
RA   Du F., Acland G.M., Ray J.;
RT   "Differential splicing of type II procollagen mRNA in canine retina.";
RL   Anim. Biotechnol. 9:15-20(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20480698; PubMed=11024291;
RA   Du F., Acland G.M., Ray J.;
RT   "Cloning and expression of type II collagen mRNA: evaluation as a
RT   candidate for canine oculo-skeletal dysplasia.";
RL   Gene 255:307-316(2000).
DR   EMBL; AF023169; AAC62178.2; -.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR000885; Fib_collagen_C.
DR   InterPro; IPR001007; VWF_C.
DR   Pfam; PF01410; COLFI; 1.
DR   Pfam; PF01391; Collagen; 18.
DR   Pfam; PF00093; vwc; 1.
DR   ProDom; PD000007; Collagen; 4.
DR   ProDom; PD002078; Fib_collagen_C; 1.
DR   SMART; SM00038; COLFI; 1.
DR   SMART; SM00214; VWC; 1.
DR   PROSITE; PS01208; VWFC; 1.
KW   Collagen.
SQ   SEQUENCE   1487 AA;  141875 MW;  25873EAC1E311DB8 CRC64;

Alignment Scores:
Pred. No.:              2.79e-13      Length:          1487
Score:                  332.50        Matches:         315
Percent Similarity:     31.42%        Conservative:    62
Best Local Similarity:  26.25%        Mismatches:      414
Query Match:            5.34%         Indels:          411
DB:                     6             Gaps:            67

US-09-759-143-110 (1-3410) x O77753 (1-1487)

Qy       3032 GTGGGGAAAGTTGGGGGTAGGGGAAAGTTGGGGGTAGGGGAAATTTTGGGCAGTGCCTTC 2973
              :::|||    :::|||    ||||||       |||                |||
Db        203 MetGlyProMetGlyProArgGlyProProGlyProAlaGlyAlaProGlyPro------ 220

Qy       2972 ATCAGCCCAGTCCTAGAGAGAGTAGAGGGGAGTGGAAGTGGGGGAACCAGGCTGGGCCA 2913
                                                |||      ||||||         |||
Db        221 ---------------------------------GlnGlyPheGlnGlyAsnProGlyGluPro 230

Qy       2912 AGAGAAGAGGGGTGGTTAGGGAAGCCGTTGAGACCTGAAGCCCCACCCTCTACCTTCCTT 2853
                 |||   |||       |||    |||:::    |||     ||||||    |||
Db        231 GlyGluProGlyValSerGly---ProMetGlyProArgGlyProProGly-ProProGl 249

Qy       2852 CAACACCCTAACCTTGGGTAACAGCATTTGGAATTATCATTTGGGATGAGTAGAATTTCC 2793

Db        249 yLysPro----------------------------------------------------- 251

Qy       2792 AAGGTCCTGGGTTAGGCATTTTGGGGGGCCAGACCCCAGGAGAAGAAGATTCTGGCAATG 2733
                                         |||   |||            :::
Db        252 -------------------------GlyAspAspGlyGluAlaGlyLys----------- 259

Qy       2732 ATCAGCCCAATGACCAGCTATCTCAGGGGACCTGATTGTTGGGGATCCCCCACCCTACCC 2673
                           |||    ||||||:::       |||    |||    |||
Db        260 ------------ProGlyLysSerGlyGluArg------GlyProProGlyPro----- 271

Qy       2672 AAATATTAGACACCAACACAGAAAAGCTAGCAATGGATTCCCTTCTACTTTGTTAAATAA 2613
                                       :::   :::   |||||||||   |||
Db        272 ---------------------GlnGlyAlaArgGlyPheProGlyThrProGlyLeuPr 284

Qy       2612 ATAAGTTAAATATTTAAATGCCTGTGTCTCTGTGATGGCAACAGAAGGACCAACAGGCCA 2553
                 ||||||                         :::           :::|||     |||
Db        284 oGlyValLysGlyHisArgGlyTyrProGlyLeuAspGlyAlaLysGlyGluAlaGlyAl 304

Qy       2552 CATCCTGATAAAAGGTAAGAGGGGGGTGGATCAGCAAAAAGACAGTGCTGTGGGCTGAGG 2493
                     :::||||||:::     |||          :::       |||         ||
Db        304 aProGlyValLysGlyGluSerGlySerProGlyGluAsnGlySerProGlyProMetGl 324

Qy       2492 GGACCTGGTTCTTGTGTGTTGCCCCTCAGGACTCTTCCCCTACAAATAAGTCATATGTTC 2433
              |                     |||    ||||||
Db        324 y--------------------ProArgGlyLeu------------------------- 328

Qy       2432 AAATCCCATGGAGGAGTGTTTCATCCTAGAAACTCCCATGCAAGAGCTACATTAAACGAA 2373
                                                 |||     :::
Db        329 ---------------------------------ProGlyGluArgGlyArgThrGlyPr 337

Qy       2372 GCTGCAGGTTAAGGGGCTTAGA------GATGGGAAACCAGGTGACTGAGTTTATTCAGC 2319
                           |||    |||       ||||||:::||||||
Db        337 oAlaGlyAlaAlaGlyAlaArgGlyAsnAspGlyGlnProGly---------------- 351

Qy       2318 TCCCAAAAACCCTTCTCTAGGTGTGTCTCAACTAGGAGGCTAGCTGTTA----------- 2270
                |||       |||          |||        ||||||
Db        352 -ProAlaGlyProProGlyProValGlyProAlaGlyGlyProGlyPheProGlyAlaPr 371

Qy       2269 ------------------------ACCCTGAGCCTGGGTAATCCACCTGCAGAGTCCCC 2235
                                       |||   :::    |||           |||:::   ||
Db        371 oGlyAlaLysGlyGluAlaGlyProThrGlyAlaArgGlyProGluGlyAlaGlnGlyPr 391

Qy       2234 GCATTCCAGTGCATGGAGCCCT---TCTGGCCTCCCTGTATAAGTCCAGACTGAAACCCC 2178
              |                  :::|||    |||          |||
Db        391 oArgGlyGluProGlyThrProGlySerProGlyProAlaGlyAlaSerGlyAsnProGl 411

Qy       2177 CTTGGAAGGCCTC---------------CAGTCAGGCAGCCCTAGAGACTGGGGAGAGAG 2133
                 :::|||:::                       :::|||:::|||          |||
Db        411 yThrAspGlyIleProGlyAlaLysGlySerAlaGlyAlaProGlyIleAlaGlyAlaPr 431

Qy       2132 G------AGAGGGACGCCCCAGCCCCCAGCTGTGCAGCTACGCACCTCA------GCAGC 2085
                        |||       ||||||       |||       |||
Db        431 oGlyPheProGlyProArgGlyProProGlyProGlnGlyAlaThrGlyProLeuGlyPr 451

Qy       2084 ACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCG 2025
                 :::|||   :::    ||||||     |||                  |||
Db        451 oLysGlyGlnThrGlyGluProGlyIleAlaGlyPheLysGlyGluGlnGlyProLysGl 471

Qy       2024 GCAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGGACCCAGT----------------- 1982
               :::|||       |||          |||       ||||||:::
Db        471 yGluProGlyProAlaGlyProGlnGlyAlaProGlyProAlaGlyGluGluGlyLysAr 491

Qy       1981 ------------------------------GAGGCAGGCCCTCCACCCCAATGTGCTGG 1953
                                                ||||||||| :::
Db        491 gGlyAlaArgGlyGluProGlyGlyAlaGlyProValGlyProProGlyGluArgGlyAl 511

Qy       1952 AAGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTC------------------- 1913
                       |||       ||||||    |||
Db        511 aProGlyAsnArgGlyPheProGlyGlnAspGlyLeuAlaGlyProLysGlyAlaProGl 531

Qy       1912 ------------------------------AAATACTACCTGTGTAGCAAAGTAAATGG 1884
                                            |||                   |||||
```

```
Db       1519 GlyIle-----------ThrGlyProSerGlyProIleGlyProProGlyProProGly 1534

Qy        178 CTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGCTTCT 119

Db       1535 LeuPro---GlyProProGly---------------ProLysGlyAlaLysGlySerSer 1548

Qy        118 CAGCCCATGCTCAACACCTGCTGCTGTGGGCACCTCAGTGGGGACACGTCTCATCACTC 59

Db       1549 GlyProThr-------------------GlyProLys-GlyGluAlaGlyHisProGl 1561

Qy         58 AGATCCTGGC 49

Db       1561 yProProGly 1564

RESULT 36
Q8VYX3
ID   Q8VYX3      PRELIMINARY;      PRT;   508 AA.
AC   Q8VYX3;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Sucrose transporter SUC2.
GN   SUC2.
OS   Brassica oleracea (Cauliflower).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Brassica.
OX   NCBI_TaxID=3712;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Coupe S.A., Sinclair B.K., Bucknell T.T., Greer L.A., Eason J.R.,
RA   Heyes J.A.;
RT   "The isolation and characterization of sucrose transporter homologs
RT   from broccoli and their role in sucrose mobilization during
RT   senescence.";
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY065840; AAL58072.1; -.
DR   InterPro; IPR003662; sub_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
SQ   SEQUENCE   508 AA;  54016 MW;  F1A2AD1D1C8EDD86 CRC64;

Alignment Scores:
Pred. No.:               2.34e-13      Length:          508
Score:                   332.50        Matches:         125
Percent Similarity:      41.57%        Conservative:    87
Best Local Similarity:   24.51%        Mismatches:      193
Query Match:             5.18%         Indels:          105
DB:                      10            Gaps:            18

US-09-759-143-110 (1-3410) x Q8VYX3 (1-508)

Qy        332 CAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATC 391

Db         26 LysIleIleSerValSerSerIleAlaAlaGlyValGlnPheGlyTrpAlaLeuGlnLeu 45

Qy        392 ACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTG 451

Db         46 SerLeuLeuThrProTyrValGlnLeuLeuGlyIleProHisLysTrpAlaSerLeuIle 65

Qy        452 CTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGAC 511

Db         66 TrpLeuCysGlyProIleSerGlyMetLeuValGlnProIleValGlyTyrHisSerAsp 85

Qy        512 CACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATC 571

Db         86 ArgCysThrSerArgPheGlyArgArgArgProPheIleValAlaGlyAlaGlyLeuVal 105

Qy        572 CTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTG------------ 619

Db        106 ThrValAlaValPheLeuIle---------GlyTyrAlaAlaAspIleGlyHisSerMet 122

Qy        620 ---------CTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTG 670

Db        123 GlyAspGlnLeuAspLysProProArgThrArgAlaIleAlaIlePheAlaLeuGlyPhe 142

Qy        671 GGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGAC 730

Db        143 TrpIleLeuAspValAlaAsnAsnThrLeuGlnGlyProCysArgAlaPheLeuAlaAsp 162

Qy        731 CTCTTCCGG---GACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATC 787

Db        163 LeuSerAlaGlyAsnAlaLysLysThrArgThrAlaAsnAlaPhePheSerPhePheMet 182

Qy        788 AGTCTTGGGGGCTGCCTGGGCTAC---------------------CTCCTGCCTGCCATT 826

Db        183 AlaValGlyAsnValLeuGlyTyrAlaAlaGlySerTyrLysAsnLeuTyrLysValVal 202

Qy        827 GACTGGGAC---ACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883

Db        203 ProPheThrMetThrLysSerCysAspLeuTyrCysAlaAsnLeuLysThrCysPhePhe 222

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943

Db        223 LeuSerIleThrLeuLeuLeuLeuValThrPheMetSerLeuCysTyrValThrGluLys 242

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003

Db        243 ---ProTrpThrProGluProThrAlaAspGly------------------LysProSer 255

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTG------GCTTTCCGGAACCTGGGCGCCCTGCTTCCC 1057

Db        256 SerValProPhePheGlyGluIlePheGlyAlaPheLysGluLeu--------------- 270

Qy       1058 CGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAG 1117

Db        271 --------------------------LysArgProMetTrpMetLeuLeuIleValThr 281

Qy       1118 CTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAG 1177

Db        282 AlaLeuAsnTrpIleAlaTrpPheProPheLeuLeuPheAspThrAspTrpMetGlyArg 301

Qy       1178 GGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGAT 1237

Db        302 GluValTyrGlyGlyAsnSerAspAlaThrAlaSerValAlaSerLysLysLeuTyrAsn 321

Qy       1238 GAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTC 1297

Db        322 AspGlyValArgAlaGlyAlaLeuGlyLeuMetLeuAsnAlaIleValLeuGlyPheMet 341

Qy       1298 TCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGT 1357

Db        342 SerLeuGlyValGluTrpValGlyArgLysMetGly---GlyAlaLysArgLeuTrpGly 360

Qy       1358 GTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTG--- 1414

Db        361 AlaValAsnPheIleLeu---------AlaIleCysLeuAlaMetThrValValValThr 377

Qy       1414 ------------------------------------------------------------ 1414

Db        378 LysGlnAlaGluAsnHisArgArgAspHisGlyGlyAlaLysThrGlyProProGlyAsn 397

Qy       1415 GTGACAGCTTCAGCC------------GCCCTCACCGGG-----------TTCACCTTC 1450

Db        398 ValThrAlaGlyAlaLeuThrLeuPheAlaValLeuGlyIleProGlnAlaIleThrPhe 417

Qy       1451 TCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAG------ 1504

Db        418 Ser------------IleProPheAlaLeuAlaSerIlePheSerThrAsnSerGlyAla 433

Qy       1505 --------------------------------CAGGTGTTCCTGCCCAAATACCGAGGG 1531

Db        434 GlyGlnGlyLeuSerLeuGlyValLeuAsnLeuAlaIleValValProGlnMetValVal 453

Qy       1532 GACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAG 1591
```

```
Qy     2175 TGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAG--------------AGGA 2131
             :::||||||     |||  |||        ||||||            :::
Db      971 ysAspGlyLeu---ProGlyHisProGlyGlnArgGlyGluThrGlyPheGlnGlyLysT 990

Qy     2130 GAGGGACGCCCCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCA 2071
             |||  |||  ||||||  |||                     ||||||   :::
Db      990 hrGlyProProGlyProProGlyVal------------ValGlyProGlnGlyProThrG 1006

Qy     2070 GAGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCC------ 2017
              |||        |||        :::   |||  ||||||  |||
Db     1006 lyGluThrGlyProMetGlyGluArgGlyHis-ProGlyProProGlyProProGlyGlu 1025

Qy     2016 ATGGGGCTAACAGGAGCGGGGAGCTGGGACCCAGTGAGGCAGGCCCTCCACCCCAATGTG 1957
              ||||||   |||                      |||:::     |||
Db     1026 GlnGlyLeuProGlyLeu-Ala---------------GlyLysGluGlyThr-------- 1037

Qy     1956 CTGGAAGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTA 1897

Db     1037 ------------------------------------------------------------ 1037

Qy     1896 GCAAAGTAAATGGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGAC 1837
                        :::  ||||||||           |||   :::
Db     1038 ------------LysGlyAspProGlyProAlaGlyLeuProGlyLysAsp-------- 1050

Qy     1836 TGGCTGAGCTGGACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAG 1777
                                                 |||||||||||     |||
Db     1051 -------------------------------------GlyProProGlyLeuArgGly-- 1057

Qy     1776 GCACTATCCAGGATGGCGAGGTCCAGGCAGATGCCCCGGC-------------------- 1737
                :::||||||  ||||||          ||||||
Db     1058 ----PheProGlyAspArgGly---------LeuProGlyProValGlyAlaLeuGlyLe 1073

Qy     1736 -------------------CCGGAACCACCCTGGCCTCGGTGGGCTCACCCACCACCACA 1696
                                 |||  ||||||  |||       |||||||||
Db     1073 uLysGlyAsnGluGlyProProGlyProProGlyPro--AlaGlySerPro--------- 1089

Qy     1695 CGTACGGAGACATCACAGGCAGAGGCCCCGCAGAGCGCGGGTGGAGGTGGGAGCAGGCCA 1636
                  |||        |||  |||            |||    |||   ||||||
Db     1090 ---GlyGluArgGlyProAlaGlyAlaAlaGlyProIleGlyIleProGly---ArgPro 1107

Qy     1635 CTGCCTCCAGCACCCACGTGTCCATTAGGG---AAGGGAGCTCCAGGC------------ 1591
               |||      |||     |||  |||  ||||||||||||||
Db     1108 GlyProGlnGlyProProGlyProAlaGlyGluLysGlyAlaProGlyGluLysGlyPro 1127

Qy     1590 TTAGGGCCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCTAGCACCTCCAGTGTCC 1531
             |||||| ||||||              |||                 ||||||   ||
Db     1128 GlnGlyPro-AlaGly------ArgAspGly------------------LeuGlnGlyPr 1139

Qy     1530 CCTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGTAGAGGGAGGCCAGTGTGTAG 1471
             |:::|||:::             ||||||  |||   |||          |||
Db     1139 oValGlyLeu------ProGlyProAlaGlyProValGly---------ProProGlyGl 1154

Qy     1470 GGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAGCTGTCACCACG 1411
                |||   :::|||     |||||||||||:::  |||                |
Db     1154 uAspGlyAsp-LysGlyGluIleGlyGluProGlyGlnLysGlySerLysGlyAspLysG 1174

Qy     1410 GCCAC---------------------------------ACTGTGGGACAGGCATGTGGCA 1384
             ||                                       :::||||||        :
Db     1174 lyGluGlnGlyProProGlyProThrGlyProGlnGlyProIleGlyGlnProGlyProS 1194

Qy     1383 CCGGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGC 1324
             :::|||:::      |||           ||||||
Db     1194 erGlyAlaAspGlyGluProGlyProArgGlyGln------------------------- 1205

Qy     1323 TGCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGC---------------- 1280
                |||   :::      |||:::   |||:::|||
Db     1206 -----GlnGlyLeuPheGlyGlnLysGlyAspGluGlyProArgGlyPheProGlyProP 1224

Qy     1279 ----------GCACTGCAGGAACAGCCCCAGGCTGCCCATCCGAACGCCTTCATCATAG 1231
                               ||||||        ||||||  |||   |||:::
Db     1224 roGlyProValGlyLeuGln--GlyLeuProGlyProProGlyGluLysGlyGluThrGl 1243

Qy     1230 TGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTACAGCCCCTCGCCC 1171
              |||         |||                            ||||||    |||  ||
Db     1243 yAspValGlyGlnMetGlyPro-----------------ProGly---ProProGlyPr 1256

Qy     1170 ACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGCTGCACAGCTCAGCC 1111
             ||||  |||                  |||      |||   :::        |||
Db     1256 oArgGlyPro-----------------SerGlyAlaProGlyAlaAspGlyProGlnGl 1270

Qy     1110 ACGAAGAGCCGGCGCAGGG---TGCGGGGCATGCGGCAGCACAGCT------GGTGCAGC 1060
                       |||   |||    ||||||  |||                  |||    ||
Db     1270 yProProGlyGlyIleGlyAsnProGlyAlaValGlyGluLysGlyGluProGlyGluAl 1290

Qy     1059 CGGGGAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGGCCCGGCATGGACAGCAGTGG 1000
             ||||||   |||   ||||||   |||   |||   ||||||        :::
Db     1290 aGlyGluProGlyLeuProGlyGluGlyGlyProProGlyProLysGlyGluArgGlyGl 1310

Qy      999 GGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCTCC 940
                                 ||||||:::                       ||||||   |||
Db     1310 uLysGlyGluSerGlyProSerGlyAlaAla-----------GlyProProGlyProLy 1326

Qy      939 TCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAG 880
             ||||||                                  |||    |||   ::
Db     1326 sGlyProPro------------------------------GlyAspAspGlyProLy 1335

Qy      879 AGGCACTCCT---------------CCTGGGTGCCCAGGTAGGGGGCCAGGGCACTGGTG 835
             :|||:::|||                  ||||||  ||||||     ||||||
Db     1335 sGlySerProGlyProValGlyPheProGlyAspProGly---ProProGlyGluProGl 1354

Qy      834 TCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCA 775
             |||:::               |||    |||  |||   |||          :::
Db     1354 yProAla---------------GlyGlnAspGlyProProGlyAsp---------LysGl 1366

Qy      774 TAGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGTCCCGGAAGAGGTCAGAGAGCAGG 718
                          ||||||   :::|||   |||||||||                    ||
Db     1366 yAspAspGlyGluProGlyGlnThrGlySerProGlyPro---------------ThrGl 1381

Qy      717 GCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGCCCAGGATG 658
             |   |||                   ||||||   |||   |||
Db     1381 yGluProGlyProSerGlyProProGlyLysArgGlyPro------------------- 1394

Qy      657 AGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGGCC 598
                          |||  |||                                  |||||||||
Db     1395 -------------Pro-Gly-----------------------------ProAlaGlyP 1400

Qy      597 CTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCAAGGACAGTGCCCAGATGAAGGGCCGG 538
             ||    :::    ::::::   ::::::    |||:::|||:::           |||||||||
Db     1400 ro---GluGlyArgGlnGlyGluLysGlyAlaLysGlyGluAlaGlyLeuGluGlyPro- 1418

Qy      537 CGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTG-------AGCCTAGGAGCGGGAC 485
                                        |||       ||||||           ||||||
Db     1419 -------------------ProGlyLysThrGlyProIleGlyProGlnGlyAlaProG 1432

Qy      484 ACAGACCAGGCCCA---GCACTGGACCAATGCCCAGCACCATGGTCATGAACTTCTCCTC 428
              :::|||||||||     |||          |||    |||                 |
Db     1432 lyLysProGlyProAspGlyLeuArgGlyIleProGlyProValGly---GluGlnGlyL 1451

Qy      427 TACCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCCAAACACACCTC 368
             |||||    |||                                   |||:::   ||||
Db     1451 euProGlySerProGly----------------------------ProAspGlyProP 1461

Qy      367 CAGGCCAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGTGCCGCAGCAGGC- 309
             |||||          |||   |||        ||||||   |||    :::
Db     1461 roGlyPro------MetGlyProProGlyLeuProGlyLeuLysGlyAspSerGlyProL 1479

Qy      308 -------------------------------------GGCTCACCCACAGCCTCTGGAC 287
                                                  |||    |||         |||
Db     1479 ysGlyGluLysGlyHisProGlyLeuIleGlyLeuIleGlyProProGlyGluGlnGly- 1498

Qy      286 CATAGTGGG-----------CCAGGCGGGTAGGGCTCAGGGGGCCGTTCAGGCACTCCA 239
                |||                ||||||      ||||||   |||          |||
Db     1499 GluLysGlyAspArgGlyLeuProGlyProGlnGlySerSerGlyProLysGlyGluGln 1518

Qy      238 GAACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAA 179
```

```
Db       1166 uLys--------------GlyGluThrGlyLeuLeuGlyAlaTyrProGlyProLysGl 1181

Qy        306 CTCACCCACAGCCTCTGGACCATAGTGGG-----------CCAGGCGGGTAGGGCTCAG 259
              |||||||          |||       |||          ||||||     |||
Db       1181 ySerProGlyValProGlyAla-LysGlyAspArgGlyValProGlyLeuSerGlyLeuP 1201

Qy        258 GGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCT 199
                ||||||  |||             :::                |||     |||
Db       1201 roGlyArgLysGlyValMetGlyAspValGlyProGlnGlyProProGlyThrAlaGlyL 1221

Qy        198 CTCCTCCTTGCTGCCGCCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCT 139
                 |||              |||         |||||||||           |||
Db       1221 euProGly--------------ProProGlyLeuProGlyAlaIleIleProGlyProL 1236

Qy        138 TTGGT-------------------------------------GCCGGTCCAGCTTCTC 118
                |||                                          ||||||
Db       1236 ysGlyAspArgGlyLeuProGlyLeuArgGlyAsnProGlyGluProGlyProProGlyP 1256

Qy        117 AGCCC 113
                |||
Db       1256 roPro 1257
```

RESULT 35
Q15094
ID   Q15094   PRELIMINARY;   PRT; 1838 AA.
AC   Q15094;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Pro-alpha-1 type V collagen.
GN   COL5A1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92105142; PubMed=1722213;
RA   Greenspan D.S., Cheng W., Hoffman G.G.;
RT   "The pro-alpha 1(V) collagen chain. Complete primary structure,
RT   distribution of expression, and comparison with the pro-alpha 1(XI)
RT   collagen chain.";
RL   J. Biol. Chem. 266:24727-24733(1991).
DR   EMBL; M76729; AAA59993.1; -.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR000885; Fib_collagen_C.
DR   InterPro; IPR001791; Laminin_G.
DR   InterPro; IPR001230; Prenyl_site.
DR   InterPro; IPR003129; TSPN.
DR   Pfam; PF01410; COLFI; 1.
DR   Pfam; PF01391; Collagen; 17.
DR   Pfam; PF02210; TSPN; 1.
DR   ProDom; PD000007; Collagen; 1.
DR   ProDom; PD002078; Fib_collagen_C; 1.
DR   SMART; SM00038; COLFI; 1.
DR   SMART; SM00282; LamG; 1.
DR   SMART; SM00210; TSPN; 1.
DR   PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
KW   Collagen.
SQ   SEQUENCE   1838 AA; 183610 MW; 5078307F6E00F0BA CRC64;

Alignment Scores:

Pred. No.:	2.68e-13	Length:	1838
Score:	333.00	Matches:	311
Percent Similarity:	31.57%	Conservative:	59
Best Local Similarity:	26.54%	Mismatches:	381
Query Match:	5.35%	Indels:	424
DB:	4	Gaps:	68

US-09-759-143-110 (1-3410) x Q15094 (1-1838)

```
Qy       3217 AAACGGCACTTAAACCCCCCCTGAGAGATAAGACCTCCCTTAGCTCAG------------ 3170
               |||   ::::::  |||     |||     ||||||     |||
Db        708 LysGlyAsnValGlyProGlnGlyGluProGlyProProGlyGlnGlnGlyAsnProGly 727

Qy       3169 GCAGGGGGTGCTCCTGAGTTTCTGTGTGAGATTCCCCAAGCACAGATATACTCTGGGGGC 3110
              |||   |||   |||               ||||||     |||        |||
Db        728 AlaGlnGlyLeuProGly--------------ProGlnGlyAlaIleGlyProProGly 742

Qy       3109 TGAGATGGACAAAGGCTTGGGAAACCGCACTTTGT------------GCTTCTGGTCCT 3063
                  |||      |||||||||||                          |||   ||||||
Db        743 GluLysGlyPro---LeuGlyLysProGlyLeuProGlyMetProGlyAlaAspGlyPro 761

Qy       3062 GCAGTAGCTCCAAACAGGGTTGTGGAGCTGGTGGGGAAAGTTGGGGGTAGGGGAAAGTTG 3003
                     |||    :::     |||     ||||||     |||   |||      |
Db        762 ProGlyHisProGlyLysGlu--GlyProProGlyGluLysGlyGlyGlnGlyProProG 781

Qy       3002 GGGGTAGGGGAAATTTTGGGCAGTGCCTTCATCAGCCCAGTCCTAGAGAGAGTAGAGGG- 2944
              ||    :::|||        |||            |||   ||||||      :::|||
Db        781 lyProGlnGlyProIleGly-----------TyrProGlyProArgGlyValLysGlyA 797

Qy       2943 -------------------------------------GAGTGGAAGTGGGGGGAACCAG 2922
                                                    :::    :::    |||   ||||
Db        797 laAspGlyIleArgGlyLeuLysGlyThrLysGlyGluLysGlyGluAspGlyPheProG 817

Qy       2921 GCTGGGCCAAGAGAAGAGGGGTGGTTAGGGAAGCCGTTGAGACCTGAAGCCCCACCCTCT 2862
              ||:::            |||:::      :::      |||        |||
Db        817 lyPheLysGlyAspMetGlyIleLysGlyAspArgGlyGluIleGlyProPro------- 834

Qy       2861 ACCTTCCTTCAACACCCTAACCTTGGGTAACAGCATTTGGAATTATCATTTGGGATGAGT 2802

Db        834 ------------------------------------------------------------ 834

Qy       2801 AGAATTTCCAAGGTCCTGGGTTAGGCATTTTGGGGGGCCAGACCCCAGGAGAAGAAGATT 2742
                         ||||||                     |||:::     |||
Db        835 -----------GlyPro---------------ArgGlyGluAspGlyProGluGlyProL 846

Qy       2741 CTGGCAATGATCAGCCCAATGA----------------------CCAGCTATCTCAGGGG 2704
                |||      ||||||                             |||        |||
Db        846 ysGlyArgGlyGlyProAsnGlyAspProGlyProLeuGlyProProGlyGluLysGlyL 866

Qy       2703 ACCTGATTGTTGGGGATCCCCCACCCTACCCAAATATTAGACACCAACACAGAAAAGCTA 2644
                |||   |||        |||   ||||||                  |||    |||   |
Db        866 ysLeuGlyValProGlyLeuProGlyTyrProGly------ArgGlnGlyProLysGlyS 884

Qy       2643 GCAATGGATTCCCTTCTACTTTGTTAAATAAATAAGTTAAATATTTAAATGCCTGTGTCT 2584
              ||   |||||||||                                   |||
Db        884 erIleGlyPheProGly-----------------PheProGlyAlaAsnGlyGluLysG 898

Qy       2583 CTGTGATGGCAACAGAAGGACCAACAGGCCACATCCTGATAAAAGGTAAGAGGGGGGTGG 2524
                       |||   |||        |||           :::|||:::||||||
Db        898 lyGlyArgGlyThrProGlyLysProGlyPro---------ArgGlyGlnArgGly---- 913

Qy       2523 ATCAGCAAAAAGACAGTGCTGTGGGCTGAGGGGACCTGGTTCTTGTGTGTTGCCCCTCAG 2464
                                                                   |||:::|
Db        914 -----------------------------------------------------ProThrG 916

Qy       2463 GACTCTTCCCCTACAAATAAGTCATATGTTCAAATCCCATGGAGGAGTGTTTCATCCTAG 2404
              ||
Db        916 ly---------------------------------------------------------- 916

Qy       2403 AAACTCCCATGCAAGAGCTACATTAAACGAAGCTGCAGGTTAAGGGGCTTAGA------G 2350
                   |||       |||                          :::|||   |||
Db        917 -----ProArgGlyGlu-----------------------ArgGlyProArgGlyIleT 927

Qy       2349 ATGGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCTC 2290
               |||||||||||                    ||||||       |||    |||
Db        927 hrGlyLysProGly-----------------ProLysGlyAsnSerGlyGlyAspGlyP 941

Qy       2289 AACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCAGAGTCCCCGCATT 2230
                 |||                                ||||||   |||          :
Db        941 roAlaGly---------------------------ProProGlyGluArgGlyProA 951

Qy       2229 CCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACCCCC------T 2176
              ::               |||:::|||   |||                      |||
Db        951 snGlyProGlnGlyProThrGlyPheProGlyProLysGlyProProGlyProProGlyL 971
```

```
Qy     2131 --------AGAGGGACGCCCCAGCCC---------CCAGCTGTGCAGCTACGCACCTCAG 2089
                    ::::|||   |||   |||         ||||||:::     |||
Db      573 roGlyAlaLysGlyProProGlyProGlnGlyGluProAlaLeuSerGlyArg---LysG 592

Qy     2088 CAGCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAG 2029
                 ||||||             |||   :::   |||             ||||
Db      592 lyAspGlnGlyProProGlyProProGlyPheProGlyProProGlyProAlaGlyProA 612

Qy     2028 CCCGGCAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGGACCCAGTGAGGCA------- 1976
            ||          |||    |||       ||||||   ||||||     |||
Db      612 laGlyProPro---GlyTyrGlyProGlnGlyGluProGlyProLysGlyAlaGlnGlyV 631

Qy     1975 --------------GGCCCTCCACCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTG 1930
                          |||||||||             |||     :::   |||
Db      631 alProGlyValLeuGlyProPro------------GlyGlu--AlaGlyLeuLysGlyGl 646

Qy     1929 GCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGG 1870
              ||||||   |||                                  ||||||   ||
Db      646 uProSerThrSer------------------------------ThrProAspLeuGl 655

Qy     1869 CCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCATA 1810
            |           |||
Db      655 yProProGlyProPro------------------------------------------ 660

Qy     1809 AACAGGGATGGGGCCACCTGGGACAGCAGGA---AGGCACTATCCAGGATGGCGAGGTCC 1753
                          ||||||||||||   ||||||   |||    ||||||     |||||
Db      661 ---------GlyProProGlyGlnAlaGlyProArgGlyLeuProGlyLeuProGlyPr 677

Qy     1752 A---GGCAGATGC------------CCCGGCCCGGAA----------------------- 1731
            |     |||:::|||                ||||||||||:::
Db      677 oValGlyLysCysAspProGlyLeuProGlyProAspGlyGluProGlyIleProGluAl 697

Qy     1730 -------------CCACCCTGGCCTCGGTGGGCTCACCCACCACCACACGTACGGAGACA 1684
                         ||||||   |||:::   |||:::                |||:::
Db      697 aGlyCysProGlyProProGlyProLys--GlyAsnGlnGlyPheProGlyThrLysGly 716

Qy     1683 TCACAGGCAGAGGCCCCGCAGAGCGCGGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCA 1624
            |||              |||         |||     |||             |||
Db      717 Ser------------ProGlyCysProGlyGluMetGlyLysProGlyArgProGlyGlu 732

Qy     1623 CCCACGTGTCCATTAGGGAAGGGAGCTCCAGGCTTAGGG--------------------- 1585
            |||      |||      ||||||   |||   :::|||
Db      733 ProGlyIleProGlyAlaLysGlyGluProSerValGlyArgProGlyLysProGlyLys 752

Qy     1585 ------------------------------------------------------------ 1585

Db      753 ProGlyPheProGlyGluArgGlyAsnAlaGlyGluAsnGlyAspIleGlyLeuProGly 772

Qy     1584 --------------------CCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCTA 1546
                                 |||||||||                           |||
Db      773 LeuProGlyLeuProGlyThrProGlyArg------------------GlyGlyLeuAsp 786

Qy     1545 GCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGTAGAGG 1486
             ||||||      |||  |||                 |||   |||||||||   :::||
Db      787 GlyProProGlyAspPro-Gly-----------GlnProGlySerProGlyAlaLysGl 802

Qy     1485 GAGGCCA---GTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCG 1429
            |   |||      |||         |||      |||   :::
Db      802 ySerProGlyArgCysIleProGlyProArgGlyThrGlnGlyLeuProGlyLeuAsnGl 822

Qy     1428 GCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACCGGCAGCCACAGGG 1369
             ||||||      |||                ||||||              |||:::||
Db      822 yLeuLysGlyGlnProGlyArgArgGlyAspThrGly--------------ProLysGl 837

Qy     1368 AAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCTGCACCAGCCGGTCC 1309
            |         |||     |||                  |||         |||     ||
Db      837 yAsp---ProGlyIleProGlyMet---------AspArgSerGlyValProGlyAspPr 853

Qy     1308 ATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACAGCCCCAGGCTGCCCATC 1249
            |   |||         ||||||                      |||||||||||||
Db      853 oGlyPro---------ProGly---------------------ThrProGlyCysProGl 863

Qy     1248 CGAACGCCTTCATCATAGTGTCTCCGGG---------------------CCTCGGTGCCC 1210
             |||                                  |||                 |||   ||
Db      863 yGluMetGlyProProGlyGlnLysGlyTyrProGlyAlaProGlyPheProGlyProPr 883

Qy     1209 GGCTCAG---------------CTCTGGGCACGCCCTGGTACAGCCCCTCGCCCACGAAA 1165
            |    :::                 :::            |||||||||     |||
Db      883 oGlyGluLysGlyGluValGlyMetMetGlyTyrProGlyThrThrGlyProProGlyLe 903

Qy     1164 TCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGCTGCACAGCTCAGCCACGAAG 1105
             |||   |||                         |||   |||         |||  :::||
Db      903 uProGlyLys---------------------ProGlySer----------GlnGlyGlnAr 913

Qy     1104 AGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACAGCTGGTGCAGCCGGGGAAGCAGGGCG 1045
            |        |||                   |||         ||||||   |||||
Db      913 gGlySerLeuGly-----------------IleProGlyMetLysGlyGluLysGlyAr 927

Qy     1044 CCCAGGTTCC--------------------GGAAAGCCAAGCGGGCCCGGCATGGACAG 1006
            |||||||:::                      |||||||||   ||||||           ::
Db      927 gProGlyAlaLysGlyGluArgGlyGluLysGlyLysPro---GlyProSerGlnThrTh 946

Qy     1005 CAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTT-------------------------- 972
            :                    :::|||         |||
Db      946 rLeuLeuLysGlyAspLysGlyGluProGlyLeuLysGlyPheValGlyAsnProGlyGl 966

Qy      971 -CTGCTGGCTCGGTGGGGCCCAGCGCTGCCT----------------------------- 942
                       |||       |||   ||||||
Db      966 uLysGlyAsnArgGlyAsnProGlyLeuProGlyProLysGlyLeuGluGlyLeuProGl 986

Qy      941 ----CCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGG 886
                |||   ||||||                        |||         |||   ||
Db      986 yLeuProGlyProPro----------------------GlyProArgGlyAspThrGl 998

Qy      885 CCAAAGAGGCACTCCTCCTGGGTGCCCAGGTAGGGGGCCAGGGCACTGGTGTCCCAGTCA 826
            |   ||||||               |||||||||         |||       |||
Db      998 ySerArgGlyAsn------------ProGlyArgProGlyProHisGlyMetProGlySe 1014

Qy      825 ATGGCAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCATAGACAGAG 766
                              |||   ||||||                              :::
Db     1014 rMetGlyIleMetGlyValProGlyPro--------------------------LysGl 1025

Qy      765 TAGGCCTGGCGACAGTGGT---CCGGGT------------CCCGGAAGAGGTCAGAGAGC 721
              |||   |||   ||||||   ||||||             |||||||||
Db     1025 yArgLysGlyThrSerGlyLeuProGlyLeuAlaGlyArgProGlyArgThrGlyIleHi 1045

Qy      720 AGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGCCCAGG 661
             ||||||       :::         ||||||:::            |||   |||   |||
Db     1045 sGlyProGlnGlyAspLysGlyGluProGlyTyrSerGluGlyAlaArgProGlyPro-- 1064

Qy      660 ATGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCG 601
                      |||   |||         |||||||||              |||      ::
Db     1065 ---------ProGlyProLysGly---AspProGly---------LeuProGlyAspLy 1077

Qy      600 GCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCA---AGGACAGTGCCCAGATGAAG 544
            :              |||||||||      |||   |||   ::::::   |||
Db     1077 sGlyLysLysGlyGluArgGlyValProGlyProGlyGlnSerGlyProAlaGlyPr 1097

Qy      543 GGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTAGGAGCGGGACA 484
               |||      |||        :::   ||||||
Db     1097 oAspGlyAlaProGlySerProGlySerProGlyHis------------------ProGl 1111

Qy      483 CAGACCAGGCCCAGCACTGGAC--------------------CAATGCCCAGCACCATG 445
             :::|||||||||||   |||                                 |||   |||
Db     1111 yLysProGlyProAlaGlyAspLeuGlyLeuLysGlyGlnLysGlyPheProGlyProPr 1131

Qy      444 GTCATGAACTTCTCCT---------------CTACCCCCACTTCCAGCAGCAGAGGCGGC 400
                      |||   |||                    ||||||   ||||||
Db     1131 oGlySerThrGlyProProGlyProProGlyLeuProGlyLeuProGlyProMetGly-- 1150

Qy      399 ACATAGGTGATGCCTGCGGCCAAACACA--------------------CCTCCAGGCCA 361
                            :::|||   :::                   |||||||||:::
Db     1151 -------------MetArgGlyAspGlnGlyArgAspGlyIleProGlyProProGlyGl 1166

Qy      360 AAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGTGCCGCAGCAGGC------GG 307
            ::::                              ||||||   ||||||     |||     ||
```

```
Qy   317 -------------GCAGCAGGC------GGCTCACCCACAGCCTCTGGACCATAGTGGG 278
                      |||  |||      |||||||||      ::::|||||| ||||||
Db   864 roGlyLeuAspGlyAlaGlnGlyGluArgGlySerProGlyLeuAsnGlyPro-SerGly 883

Qy   277 CCA-----------GGCGGGTAGGGCTCAGGGGGCCGTTCAGGCACTCCAGAACTGCTT 230
         |||           |||       |||      |||   :::|||  |||
Db   884 ProProGlyProValGlyProGlnGlyGluArgGlyAlaAsnGlyPhePro--------- 900

Qy   229 CGTCTCGGCTCTGCTCCAGAAGCT---GCGGCCTCTCCTCCTTGCTGCCGCCAACTGCCT 173
               ||||||     ||||||  ||||||   ||||||     :::|||
Db   901 ------GlySerGlnGlyGluAlaGlyAlaAlaGlyProPro-GlySerAlaGlyGluPr 918

Qy   172 AGGAATC-----------------AGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCC 131
            |||:::                            ||||||     |||   |||       ||
Db   918 oGlyLeuArgGlyAspAsnGlyAsnAspGlyAlaProGlyGlnAlaGlyProProGlyPr 938

Qy   130 GGTCCAGCTTCTCAGCCCATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTGGGGACAC 71
          |             |||      |||                  |||:::||
Db   938 oThr---------GlyProAlaGlyTyrPro--------------------GlyGluTh 948

Qy    70 GTCTCATCACTCAGATCCTGGCCGAGGCGCGCGGCTGTCACCC-----------GGAGC 23
          |              ||||||:::        |||   |||              |||||
Db   948 rGlyGlnProGlySerProGlyLysAspGlyProLeuGlyProValGlyArgSerGlyAl 968

Qy    22 CAGCGCGTGCAGG 10
          |          |||
Db   968 aLysGlyAlaArg 972
```

```
RESULT 34
Q9QZS0
ID   Q9QZS0     PRELIMINARY;     PRT;  1669 AA.
AC   Q9QZS0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Alpha 3 collagen IV.
GN   COL4A3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KIDNEY;
RX   MEDLINE=20005934; PubMed=10534397;
RA   Lu W., Phillips C.L., Killen P.D., Hlaing T., Harrison W.R.,
RA   Elder F.F.B., Miner J.H., Overbeek P.A., Meisler M.H.;
RT   "Insertional mutation of the collagen genes col4a3 and col4a4 in a
RT   mouse model of alport syndrome.";
RL   Genomics 61:113-124(1999).
DR   EMBL; AF169387; AAD50449.1; -.
DR   MGD; MGI:104688; Col4a3.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR001442; ProcollagnC4.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   Pfam; PF01413; C4; 2.
DR   Pfam; PF01391; Collagen; 21.
DR   ProDom; PD000007; Collagen; 6.
DR   ProDom; PD003923; ProcollagnC4; 1.
DR   SMART; SM00111; C4; 2.
DR   PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
KW   Collagen.
SQ   SEQUENCE   1669 AA;  161769 MW;  30976E59739A47B2 CRC64;

Alignment Scores:
Pred. No.:               2.64e-13     Length:         1669
Score:                   333.00       Matches:        299
Percent Similarity:      29.65%       Conservative:   53
Best Local Similarity:   25.19%       Mismatches:     379
Query Match:             5.35%        Indels:         457
DB:                      11           Gaps:           68

US-09-759-143-110 (1-3410) x Q9QZS0 (1-1669)
```

```
Qy  3029 GGGAAAGTTGGGGGTAGGGGAAAGTTGGGGGTA-----------GGGGAAATTTTGGGC 2982
         |||      |||   |||||:::   |||:::                  |||:::|||
Db   318 GlyPheProGlyLeuArgGlyGluAlaGlyIleLysGlyArgLysGlyAspIleGlyPro 337

Qy  2981 AGTGCCTTCATCAGCCCAGTC-----------------CTAGAGAGAGTAGAGGGGAGT 2940
             |||     |||                          ||||||:::
Db   338 ProGlyPheProGlyProThrGluTyrTyrAspAlaTyrLeuGluLys----------- 353

Qy  2939 GGAAGTGGGGGGAACCAGGCTGGGCCAAGAGAAGAGGGGTGGTTAGGGAAGCCGTTGAGA 2880
         |||      |||               |||      :::|||      |||   |||
Db   354 GlyGluArgGlyMetProGlyLeuProGlyProLysGlyAlaArgGly---ProGlnGly 372

Qy  2879 CCTGAAGCCCCACCCTCTACCTTCCTTCAACACCCTAACCTTGGGTAACAGCATTTGGAA 2820
         |||      ||||||      :::|||   :::|||
Db   373 ProSerGlyProProGly--ValProGlySerPro------------------------ 383

Qy  2819 TTATCATTTGGGATGAGTAGAATTTCCAAGGTCCTGGGTTA------------------ 2779
                                           |||||||||
Db   384 -------------------GlyLeuSerArgProGlyLeuArgGlyProIleGlyTrpP 397

Qy  2778 --GGCATTTTGGGGGGCCAG-----------ACCCCAGGAGAAGAAGATTCTGGCAATG 2733
            |||:::   |||   :::                 ||||||::::::       |||
Db   397 roGlyLeuLysGlySerLysGlyGluArgGlyProProGlyLysAspThrValGlyProP 417

Qy  2732 ATCAGCCCAATGA-------------------------------------CCAGCTA 2713
            |||                                                |||
Db   417 roGlyProLeuGlyCysProGlySerProGlyProProGlyProProGlyProProGlyC 437

Qy  2712 TCTCAGGGGACCTGATTGTT---------------------------GGGGAT---- 2687
            ||||||::::::                                  ||||||
Db   437 ysProGlyAspIleValPheLysCysSerProGlyGluHisGlyMetProGlyAspThrG 457

Qy  2686 --CCCCCACCCTACCCAAATATTAGACACCAACACAGAAAAGCTAGCAATGGATTCCCTT 2629
           ||||||      |||   :::                   |||      |||   |||
Db   457 lyProProGlyValProGlyLeu------AspGlyProLysGlyGluProGlySerPro- 474

Qy  2628 CTACTTTGTTAAATAAATAAGTTAAATATTTAAATGCCTGTGTCTCTGTGATGGCAACAG 2569
                                                  |||
Db   475 --------------------------------------CysThrGluCysHisCysPheP 482

Qy  2568 AAGGACCAACAGGCCACATCCTGATAAAAGGTAAGAGGGGGGTGGATCAGCAAAAAGACA 2509
            ||||||   |||               :::   |||      |||:::|||      |||
Db   482 roGlyProProGly---------ValProGlyPheProGlyLeuAspGlyIleLysGlyI 499

Qy  2508 GTGCTGTGGGCTGAGGGGACCTGGTTCTTGTGTGTTGCCCCTCAGGACTCTTCCCCTACA 2449
                |||   |||
Db   499 leProGlyGlyArgGly------------------------------------------ 504

Qy  2448 AATAAGTCATATGTTCAAATCCCATGGAGGAGTGTTTCATCCTAGAAACTCCCATGCAAG 2389

Db   504 ------------------------------------------------------------ 504

Qy  2388 AGCTACATTAAACGAAGCTGCAGGTTAAGGGGCTTAGAGATGGGAAACCAGGTGACTGAG 2329
                                |||   ||||||:::   |||   ||||||
Db   505 ---------------------ValProGlyLeuLys---GlyAsnProGly------- 513

Qy  2328 TTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCTCAACTAGGAGGCTAGCTGTTAA 2269
                                |||   |||   :::   |||   |||
Db   514 ---------------------SerProGlySerAlaGlyLeuProGly---------- 522

Qy  2268 CCCTGAGCCTGGGTAATCCACCTGCAGAGTCCCCGCATTCCAGTGCATGGAGCCCTTCTG 2209
                         |||   :::      |||                           |
Db   523 --------PheAlaGlyPheProGlyAspGlnGlyHisProGlyLeuLysGlyAspLysG 540

Qy  2208 GCCTCCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCTA 2149
         ||                     :::|||||||||         |||   |||:::|||
Db   540 ly---------------------AspThrProLeuProTrpGlyGlnValGlyAsnProG 553

Qy  2148 GAGACTGGGGAGAGAGG------------------------------------------ 2132
           |||   |||   |||
Db   553 lyAspProGlyLeuArgGlyLeuProGlyArgLysGlyPheAspGlyThrProGlyGlyP 573
```

```
Db    258 gGlyGlu---------AspGlyMet---AlaGlyAsnProGlySerValGlyPro---Il 273

Qy   2129 AGGGACGCCCCAGCCCCCAGCTGTGCAG-----------CTACGCACCTCAGCAGCACA 2082
            |||   |||   ||||||                          :::    |||  ||
Db    273 eGlyProProGlyProProGlyPheProGlySerSerGlyAlaLysGlyAspAlaGlyGl 293

Qy   2081 GGGTGGCAGCAGAGAGCCA---CATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCG 2025
          |  |||  |||  |||          |||:::     |||      |||
Db    293 nSerGlyAspArgGlyProMetGlySerSerGlyAlaProGlyAsnAsnGlyAlaAlaGl 313

Qy   2024 GCAGCCCCATGGGGCTAACAGG---------------------------------AG 2001
            ||||||     |||
Db    313 yGlnProGlyGlnProGlyArgProGlySerProGlyAspAspGlyAsnProGlyThrLy 333

Qy   2000 CGGGGAGCTGGGACCCAGTGAGGCAGGCCCT--------------------------- 1970
            ||||||:::||||||:::    ||||||
Db    333 sGlyGluMetGlyProAla---GlyGlyProGlySerProGlyPheProGlyProGlnGl 352

Qy   1969 -CCACCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAA 1911
            ||||||     |||||||:::
Db    352 yProProGlyAlaAlaGlyAla------------------------------------ 359

Qy   1910 ATACTACCTGTGTAGCAAAGTAAATGGCGACCA----GACCCAGGCCTGCGGCAGACACC 1855
                                       |||:::    |||   |||       :::||
Db    360 ------------GlnGlyAlaAlaGlyAsnLysGlyAspGlnGlyArgAspGlySerPr 375

Qy   1854 ATATAGGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCATAAACAGGGATGGGGCC 1795
          |     |||   :::   |||   |||||||||                    |||
Db    375 oGlySerGlnGlyGluValGlyAlaAlaGlyGlnProGlyAlaGlnGlyGluArgGlyLe 395

Qy   1794 ACCTGGGACAGCAGGAAGGCACTATCCAGGATGGCGAGGTCCA---------GGCAGATG 1744
            ||||||     |||:::        |||   |||||||||         |||
Db    395 uProGlyLeuProGlyGlnGlu------GlyLysArgGlyProLeuGlyAlaGlyGlyPr 413

Qy   1743 CCCCGGCCCGGAACCACCCTG-GCCTCGGTGGGCTCACCCACCACC------------- 1699
            |||||||||      |||          |||   |||   |||
Db    413 oProGlyProProGlyProSerGlyGluArgGlyLeuProGlyThrProGlyMetSerGl 433

Qy   1698 ------------------------------------------------------ACACG 1694
                                                                   ||
Db    433 yArgProGlyAlaAlaGlyValAlaGlyLysAspGlyAlaArgGlyLeuGlnGlyGluAr 453

Qy   1693 TACGGAGACATCACAG---------GCAGAGGCCCCGCAGAGCGCGGGTGGAGGTGGGAG 1643
          |   |||     |||          |||   ||||||   |||||||||   |||::
Db    453 gGlyGluAlaGlyGlnGlnGlyValAlaGlyAlaProGlySerAlaGlyGlnThrGlyAl 473

Qy   1642 CAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGGGAAGGGAGCTCCAGGC---TTAGG 1586
          :|||           ||||||    |||          |||||||||||       ||
Db    473 aArg---------GlyAlaProGlyPheProGlyAlaAspGlyAlaProGlyGluArgGl 490

Qy   1585 GCCTGGCAGGAAGC---TGGTCATCAGGCTGTCCTCACTGCTAGCACCTCCAGTGTCCCC 1529
            |||| ||||||:::                |||    |||         |||   |||
Db    490 yPro-AlaGlyAsnProGlyValAspGlyAsnPro---------GlyGluGlnGlyProG 507

Qy   1528 TCGGTATTTGGGCAGGAACACCT-------------------------------GCTT 1502
            ||||||       |||   |||
Db    507 lnGlyIle---GlnGlyProProGlyLeuGlnGlyProSerGlyAsnLysGlyAspThrG 526

Qy   1501 CTCCCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAA 1442
            |||   |||:::          |||    |||   |||:::   |||
Db    526 lyProSerGlyLysAspGlyAspValGlyArgProGlyAlaValGly------------ 541

Qy   1441 CCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACC 1382
                                     ||||||                         |
Db    542 -------------------------ProArg--------------------GlyGluA 546

Qy   1381 GGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCTG 1322
           ||   |||   |||                             |||     :::
Db    546 rgGlyProAlaGly----------------------------GluAlaGlyLeuGluG 556

Qy   1321 CACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACAGCCC 1262
            |||   |||                 ||||||                |||   |||
Db    556 lyProProGly---------------SerProGlyAlaGluGlyGluIleGlyAlaAla- 570

Qy   1261 CAGGCTGCCCATCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGC 1202
            |||   |||   |||:::          :::   ||||||      |||
Db    571 --GlyArgProGlyGluGlnGlyPheGlnGlyLeuProGlyProAlaGlyPro------- 587

Qy   1201 TCTGGGCACGCCCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGT 1142
                     ||||||   |||   ||||||      |||                  |
Db    588 -----------ProGlyGluAlaGlyArgPro---GlyProValGly---------GluS 600

Qy   1141 CATGAGTGCCATCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCAT 1082
          ||    :::|||         |||         |||         |||
Db    600 erGlyIleProGlyGluProGlyAlaSerGlyGluArgGlyGluArgGly---------A 617

Qy   1081 GCGGCAGCACAGCTGGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGCG 1022
                |||   :::|||         |||  ||||||         ||||||
Db    617 laProGlyGluValGlyValSerGlySerArgGlyAlaProGlyGluArgGlyProSerG 637

Qy   1021 GGCCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTT---------- 972
          |||||      :::         |||         ||||||
Db    637 lyProSerGlyGluValGlyGlnAlaGlyProProGlyProAlaGlyAlaArgGlyAspA 657

Qy    971 -------------CTGCTGGCTCGGTGGGGCCCAGCGCTGCCT---------------C 941
                       |||:::            |||         |||               |
Db    657 laGlyAlaGlnGlyLeuIleGlyMetProGlyGluArgGlyProIleGlyArgAsnGlyP 677

Qy    940 CTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAA 881
          |||||                           |||      |||      |||:::|
Db    677 roGln-----------------------------GlyAsnArgGlyLeuThrGlyGluA 687

Qy    880 GAGGCACTCCTCCTGGGTGCCCAGGTAGGGGGCCAGGGCACTGGTGTCCCAGTCAATGGC 821
          |||||      |||   |||||||||   ||||||
Db    687 rgGly---GlnAspGlyGluProGlyArg---ProGlyGlu------------------ 698

Qy    820 AGGCAGGAGGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGC 761
                |||:::||||||||||||                          :::|
Db    699 -----AlaGlyAlaProGlySerPro--------------------------GlyGlnP 708

Qy    760 CTGGCGACAGTGGTCCGGGTCCCGGAAGAGGTCAGAGA------GCAGGGCCTCCAGTGG 707
          |||||  ||||||         :::|||   |||      |||||||||      |
Db    708 roGlyProSerGlyLeuValSerAlaLysGlyAspArgGlyGluAlaGlyProAlaGlyG 728

Qy    706 AGTGAAGCACACCT---------GGCCACAGAAGTCCAGCAGCCCCACGCCCAGGATGAG 656
          ||         |||          |||   |||   ||||||
Db    728 luProGlyProProGlySerSerGlyGlnArgGlyProAlaGly--------------- 742

Qy    655 CAGTGCCAGCTCCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGGCCCT 596
                   |||  |||||||||            |||   |||           :::||||||
Db    743 --------Ala-GlnGlyPro-----------GlnGlyProThrGlyLeuSerGlyPro 755

Qy    595 TGG--------------------------------GATGAGAAAGAGGCTCAGCAGGAT 569
                                                |||         :::
Db    756 ThrGlyGluMetGlyGlnThrGlySerAspGlyLysAspGlyAlaLysGlyAspThrGly 775

Qy    568 GCCCAAGGACAGTGCCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTC 509
          |||:::|||         :::   ||||||         |||
Db    776 AlaArgGlyTyrProGlyGluAlaGlyProIleGlyAlaProGly--------------- 790

Qy    508 ACTGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCAC 449
                  |||   |||   |||         :::                  :::|||   :::|||
Db    791 ---Asn-GluGlyArgGluGlyArgLysGlySerArgGlySerGlyGlyIleProGlyAs 809

Qy    448 CATGGTCATGAACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGAT 389
                          |||   ||||||         |||   |||
Db    809 n---------------SerGlyThrProGlyAspProGlyArgAlaGlyProProGlySe 824

Qy    388 GCCTGCGGCCAAACACACCTCCAGGCCAAAGGTTAGCA--------------------- 351
            |||   |||:::   |||||||||         |||
Db    824 rProGlyAlaGln-GlyProProGlyProSerGlyAlaThrGlyLeuSerGlyAspGlyG 844

Qy    350 ----------------GGTTGACCAGCAAGAGCTGGGCTTTCCGGTGCC---------- 318
                          |||   |||   |||:::|||      ||||||
Db    844 lyGluArgGlyGluThrGlyProProGlyArgSerGlyGluProGlyAlaProGlyMetP 864
```

```
Db         330 rAspMetGlyValArgAlaGlyAlaLeuGlyLeuMetLeuAsnSerValValLeuGlyAl 350

Qy        1294 CTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGC 1353
                    ||||||   :::|||   |||   :::        |||               |||
Db         350 aThrSerLeuGlyValAspIleLeuAlaArgGlyValGly---GlyValLysArgLeuTr 369

Qy        1354 CAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGT 1413
                    :::      |||    :::           |||   ||||||      :::|||   ||
Db         369 pGlyIleValAsnPheLeuLeu---------AlaIleCysLeuGlyLeuThrValLeuVa 386

Qy        1414 GGTG-------------------------------------------------------- 1417
               |
Db         386 lThrLysLeuAlaGlnHisSerArgGlnTyrAlaProGlyThrGlyGlyLeuGlnAspPr 406

Qy        1418 ------------------ACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCT 1458
                                    :::   ||||||||   |||:::             ::
Db         406 oLeuProProSerGlyGlyIleLysAlaGlyAlaLeuThrLeuPheSerValLeuGlyIl 426

Qy        1459 GCAGATC---------------CTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAA 1503
               :    :::                :::|||:::   |||||||||::::::
Db         426 eProLeuAlaIleThrTyrSerIleProPheAlaLeuAlaSerIlePhe----------- 442

Qy        1504 GCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCT 1563
                                                |||    ||||||              |||||
Db         443 ------------------------SerSerThrSerGlyAlaGlyGlnGlyLeuSerLe 454

Qy        1564 G----------------------------------ATGACCAGCTTCCTGCCAGGCCC 1587
               |                                  |||    |||    |||||
Db         454 uGlyValLeuAsnLeuAlaIleValIleProGlnMetPheValSerValLeuSerGlyPr 474

Qy        1588 TAAGCCTGGA--------------------GCTCCCTTCCCTAATGGACACGTGGGTGC 1626
               |                                  |||     |||    |||    ||
Db         474 oTrpAspAlaLeuPheGlyGlyGlyAsnLeuProAlaPheValValGlyAlaValAlaAl 494

Qy        1627 TGGAGGCAGTGGC--------------CTGCTCCCACCTCCACCC 1657
               |      ||||||                |||||||||    ||||||
Db         494 aLeuAlaSerGlyIleLeuSerMetIleLeuLeuProSerProPro 509
```

RESULT 33
O97405

```
ID   O97405       PRELIMINARY;      PRT;  1378 AA.
AC   O97405;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Collagen pro alpha-chain precursor.
GN   HDCOL 1 ALPHA.
OS   Haliotis discus (Abalone).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Archaeogastropoda;
OC   Haliotidae; Haliotis.
OX   NCBI_TaxID=36094;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=MUSCLE;
RX   MEDLINE=99234051; PubMed=10215888;
RA   Yoneda C., Hirayama Y., Nakaya M., Matsubara Y., Irie S., Hatae K.,
RA   Watabe S.;
RT   "The occurrence of two types of collagen proalpha-chain in the abalone
RT   Haliotis discus muscle.";
RL   Eur. J. Biochem. 261:714-721(1999).
DR   EMBL; AB017600; BAA75668.1; -.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR000885; Fib_collagen_C.
DR   InterPro; IPR001007; VWF_C.
DR   Pfam; PF01410; COLFI; 1.
DR   Pfam; PF01391; Collagen; 17.
DR   ProDom; PD000007; Collagen; 1.
DR   ProDom; PD002078; Fib_collagen_C; 1.
DR   SMART; SM00038; COLFI; 1.
DR   SMART; SM00214; VWC; 1.
KW   Signal.
FT   SIGNAL        1     18       POTENTIAL.
SQ   SEQUENCE   1378 AA;  132583 MW;  6EEE34A65B52E634 CRC64;
```

```
Alignment Scores:
Pred. No.:               2.56e-13      Length:          1378
Score:                   333.00        Matches:         299
Percent Similarity:      30.71%        Conservative:    60
Best Local Similarity:   25.58%        Mismatches:      389
Query Match:             5.35%         Indels:          423
DB:                      5             Gaps:            66

US-09-759-143-110 (1-3410) x O97405 (1-1378)

Qy        3077 TGTGCTTCTGGTCCTGCAGTAGCTCCAAACAGGGTTGTGGAGCTGGTGGGGAAAGTTGGG 3018
               |||||||||   |||        |||   |||        |||
Db          84 CysAlaSerCysProGluLysAlaSerAsn------ThrGluGlnAlaGlnAsnThr--- 100

Qy        3017 GGTAGGGGAAAGTTGGGGGTAGGGGAAATTTTGGGCAGTGCCTTCATCAGCCCAGTCCTA 2958
                  :::|||   :::                              :::   ||||||
Db         101 ---GlnGlyAsnValLysAlaAlaProAlaTyrSerProGlnTyrTyrSerProGlyAla 119

Qy        2957 GAGAGAGTAGAGGGGAGTGGAAGTGGGGGGAACCAGGCTGGGCCAAGAGAA---GAGGGG 2901
                        |||                       |||   ||||||:::      :::|||
Db         120 ProGlyProArgGlyProProGlyProProGlyGlnProGlyProGlnGlyPheGlnGly 139

Qy        2900 TGGTTAGGGAAGCCGTTGAGACCTGAAGCCCCACCCTCTACCTTCCTTCAACACCCTAAC 2841
                  |||:::         |||       |||   ||||||
Db         140 MetArgGlyGluSerGlyAspProGlyProProGlySerThr-Gly------------- 154

Qy        2840 CTTGGGTAACAGCATTTGGAATTATCATTTGGGATGAGTAGAATTTCCAAGGTCCTGGGT 2781

Db         155 ---------------------------------------------------------Gl 155

Qy        2780 TAGGCATTTTGGGGGGCCAGACCCCAGGAGAAGAAGATTCTGGCAATGATCAGCCCAATG 2721
                  |||   |||   ||||||
Db         155 nArgGlyPheAlaGlyProProGlySer------------------------------- 164

Qy        2720 ACCAGCTATCTCAGGGGACCTG---ATTGTTGGGGATCCCCCACCCTACCCAAATATTAG 2664
                  |||   :::   |||:::      :::   |||||||||   |||         :::
Db         165 -ProGlyLeuAspGlyGluGluGlyMetProGlyAspProGlyProAlaGlyProLeuGl 184

Qy        2663 ACACCAACACAGAAAAGCTAGCAATGGATTCCCTTCTACTTTGTTAAATAAATAAGTTAA 2604
                                               |||   |||
Db         184 ySerProGlySerSerGlyArgProGlyMetPro------------------------- 195

Qy        2603 ATATTTAAATGCCTGTGTCTCTGTGATGGCAACAGAAGGACCAACAGGCCACATCCTGAT 2544
                                                       ||||||   ||||||
Db         196 --------------------------GlyMetProGlyProLysGlyHisArgGlyPr 206

Qy        2543 AAAAGGTAAGAGGGGGGTGGATCAGCAAAAAGACAGTGCTGTGGGCTGAGGGGACCTGGT 2484
                  |||:::   |||   |||   :::                        ||||||
Db         206 oSerGlyArgProGlyArgAspGlyGluAlaGlyLysGln---------GlyAsp----- 221

Qy        2483 TCTTGTGTGTTGCCCCTCAGGACTCTTCCCCTACAAATAAGTCATATGTTCAAATCCCAT 2424

Db         221 ------------------------------------------------------------ 221

Qy        2423 GGAGGAGTGTTTCATCCTAGAAACTCCCATGCAAGAGCTACATTAAACGAAGCTGCAGGT 2364

Db         221 ------------------------------------------------------------ 221

Qy        2363 TAAGGGGCTT------AGAGATGGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAA 2310
                  ||||||:::         |||   ||||||               |||   |||:::
Db         222 -LysGlyIleProGlyAlaAlaGlyProProGlySer---------SerGlyProGlnGl 238

Qy        2309 CCCTTCTCTAGGTGTGTCTCAACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCC 2250
                  |||                                              |||:::||
Db         238 yProArg---------------------------------------------GlySerPr 243

Qy        2249 ACCTGCAGAGTCCCCGCATTCCAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCA 2190
               |                  |||      ||||||                     :::::
Db         243 oGlyGluArgGlyArgAspGlySerProGlySerProGly--------------IleAr 258

Qy        2189 GACTGAAACCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAG 2130
               :   |||            :::|||:::   :::|||:::|||         |||
```

```
                 ::::::  |||  ::::::  |||     |||     |||     :::
Db      351 GlyValGluPheLeuGlyLysLysIleGly---GlyAlaLysArgLeuTrpGlyIleLeu 369

Qy     1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTG------ 1417
              |||  :::            |||  ||||||:::  ::::::  |||
Db      370 AsnPheValLeu---------AlaIleCysLeuAlaMetThrIleLeuValThrLysMet 386

Qy     1418 --------------------ACAGCTTCAGCCGCCCTCACCGGGTTCACC--------- 1447
                                     :::     |||  |||  |||
Db      387 AlaGluLysSerArgGlnHisAspProAlaGlyThrLeuMetGlyProThrProGlyVal 406

Qy     1448 --------------------TTCTCAGCCCTGCAGATC--------------------- 1465
                                 |||:::||||||  |||
Db      407 LysIleGlyAlaLeuLeuLeuPheAlaAlaLeuGlyIleProLeuAlaAlaThrPheSer 426

Qy     1466 CTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAG-------------------- 1504
            :::|||:::  |||||||||::::::            :::
Db      427 IleProPheAlaLeuAlaSerIlePheSerSerAsnArgGlySerGlyGlnGlyLeuSer 446

Qy     1505 -----------------CAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCT 1546
                                :::  :::|||:::                    |||
Db      447 LeuGlyValLeuAsnLeuAlaIleValValProGlnMetLeuValSerLeuValGlyGly 466

Qy     1547 AGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTC 1606
                  :::|||                ||||||||                   |||
Db      467 ProTrpAspAspLeuPheGlyGlyGlyAsnLeuProGly------------------Phe 480

Qy     1607 CCTAATGGACACGTGGGTGCTGGAGGCAGTGGC--------------CTGCTCCCACCT 1651
                  |||  |||  |||   |||                      :::||||||
Db      481 ValValGlyAlaValAlaAlaAlaAlaSerAlaValLeuAlaLeuThrMetLeuProSer 500

Qy     1652 CCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGT 1708
              ||||||||                    |||           |||  ::::|||
Db      501 ProProAla--------------------AspAlaLysProAlaValAlaMetGly 512
```

```
RESULT 32
Q9XHL6
ID   Q9XHL6      PRELIMINARY;      PRT;    524 AA.
AC   Q9XHL6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Sucrose transport protein SUT1.
OS   Pisum sativum (Garden pea).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Vicieae; Pisum.
OX   NCBI_TaxID=3888;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99291546; PubMed=10363367;
RA   Tegeder M., Wang X.D., Frommer W.B., Offler C.E., Patrick J.W.;
RT   "Sucrose transport into developing seeds of Pisum sativum L.";
RL   Plant J. 18:151-161(1999).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR   EMBL; AF109922; AAD41024.1; -.
DR   InterPro; IPR003662; sub_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   TIGRFAMs; TIGR01301; GPH_sucrose; 1.
KW   Transmembrane.
SQ   SEQUENCE   524 AA;  55279 MW;  B5CE3F880D2C1E6B CRC64;

Alignment Scores:
Pred. No.:              2.18e-13       Length:         524
Score:                  333.00         Matches:        140
Percent Similarity:     39.39%         Conservative:   79
Best Local Similarity:  25.18%         Mismatches:     200
Query Match:            5.19%          Indels:         137
DB:                     10             Gaps:           20

US-09-759-143-110 (1-3410) x Q9XHL6 (1-524)

Qy      256 CCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCT 315
            |||||||||  |||:::      :::                :::  :::|||  :::
Db        3 ProLeuSerSerThrLysGlnIleAsnAsnAsnAsnAsnAsnLeuAlaLysProSerSer 22

Qy      316 GCGGCACCGGAAAGCCCAG--------------------------CTCTTGCTGGTCAA 348
              |||  :::  |||                               :::::::::|||
Db       23 Leu-HisValGluThrGlnProLeuGluProSerProLeuArgLysIleMetValValAl 42

Qy      349 CCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCT 408
              :::        |||:::::        |||       ::::::  :::  |||
Db       42 aSerIleAlaAlaGlyValGlnPheGlyTrpAlaLeuGlnLeuSerLeuLeuThrProTy 62

Qy      409 GCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGT 468
              :::     :::|||:::         :::        :::         ||||||::
Db       62 rValGlnLeuLeuGlyIleHisHisThrTrpAlaAlaTyrIleTrpLeuCysGlyProIl 82

Qy      469 GCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTA 528
            :   |||::::::      |||::::::|||        ||||||          |||::
Db       82 eSerGlyMetLeuValGlnProValValGlyTyrHisSerAspArgCysThrSerArgPh 102

Qy      529 TGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCT 588
            :|||||||||||||||||||    |||  |||:::   :::   :::::::::||||
Db      102 eGlyArgArgArgProPheIleAlaAlaGlySerIleAlaValAlaIleAlaValPheLe 122

Qy      589 CATCCCAAGGGCCGGCTGGCTAGCAGGGCTG---------------------CTGTGCCC 627
            ||||         |||::: |||  |||                              |||
Db      122 uIle---------GlyTyrAlaAlaAspLeuGlyHisSerPheGlyAspAsnLeuAspLy 139

Qy      628 GGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTG 687
                  ||||||    :::   :::   ::::::|||      :::||||||
Db      139 sLysValArgProArgAlaIleGlyIlePheValValGlyPheTrpIleLeuAspValAl 159

Qy      688 TGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGG---GACCC 744
                :::        |||     ||||||||    ||||||          :::
Db      159 aAsnAsnMetLeuGlnGlyProCysArgAlaLeuLeuGlyAspLeuCysAlaGlyAsnGl 179

Qy      745 GGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCT 804
                  |||  |||   :::   ::::::|||   ::::::::|||          ||
Db      179 nArgLysThrArgAsnAlaAsnAlaPhePheSerPhePheMetAlaValGlyAsnValLe 199

Qy      805 GGGCTACCTCCTGCCTGCCATTGAC-----------------------TGGGACACCAG 840
            |||||||       |||                                  |||
Db      199 uGlyTyrAlaAlaGlyAlaTyrSerLysLeuTyrHisValPheProPheThrLysThrGl 219

Qy      841 TGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCAT 900
              |||     |||           :::  |||  |||      :::  |||::
Db      219 uAlaCysAsnValTyrCysAlaAsnLeuLysSerCysPhePheLeuSerIleAlaLeuLe 239

Qy      901 CTTCCTCACCTGCGTAGCAGCC------------------ACACTGCTGGTGGCTGAGGA 942
            :  :::          ||||||                  |||   |||:::||||||::
Db      239 uThrValLeuAlaThrAlaAlaLeuIleTyrValLysGluThrProLeuIleAlaGluLy 259

Qy      943 GGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCA 1002
            :|||  :::          :::|||  |||          :::
Db      259 sAlaValVal---------ThrAlaGluAspGlyGlySerAsnGlyGlyMet-------- 273

Qy     1003 CTGCTGTCCATGCCGGGCCCGCTTG------GCTTTCCGGAACCTGGGCGCCCTGCTTCC 1056
                   ||||||  ::::|||  |||||||:::  |||
Db      274 -------ProCysPheGlyGlnLeuSerGlyAlaPheLysGluLeu-------------- 286

Qy     1057 CCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGA 1116
                                          |||  :::      |||  :::
Db      287 --------------------------LysArgProMetTrpIleLeuLeuLeuValTh 297

Qy     1117 GCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGA 1176
                 :::|||:::|||          |||  ||||||  ||||||::::::|||::
Db      297 rCysLeuAsnTrpIleAlaTrpPheProPheLeuLeuPheAspThrAspTrpMetGlyLy 317

Qy     1177 GGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACAC---TA 1233
            :    ::::|||                    ||||||        |||   ||
Db      317 sGluValTyr--------------------GlyGlyThrValGlyGluGlyHisAlaTy 330

Qy     1234 TGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGT 1293
            ||||    |||||||||   |||::::|||||||||    |||                :::
```

```
Qy       1475 ACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGAC 1534
                             |||:::   :::
Db        384 --------------GlnHisGlnArgGlnHis-------------------------- 389

Qy       1535 ACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCT 1594

Db        389 ------------------------------------------------------------ 389

Qy       1595 GGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCA 1654
                                                ||| :::|||  |||  ||||||
Db        390 --------------------------------GlyAlaAsnGlyGlnLeuLeuProPro 398

Qy       1655 CCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCC 1714
                |||      |||     |||     :::|||:::        ::::::|||  |||
Db        399 SerAla------GlyValLysAlaGlyAlaLeuSerLeuPheSerIleLeuGlyIlePro 416

Qy       1715 ACCGAGGCCAGGGTG---GTTCCG--------------------------------GGC 1738
                                :::|||                                 |||
Db        417 LeuSerIleThrPheSerIleProPheAlaLeuAlaSerIleTyrSerSerGlySerGly 436

Qy       1739 CGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCC 1798
                 |||      |||   |||   :::|||:::   |||    ::::::   |||:::
Db        437 AlaGlyGlnGlyLeuSerLeuGlyValLeuAsnLeuAlaIleValValProGlnMetIle 456

Qy       1799 CCATCCCTGTTTATGGGCTCCATTGTCCAGCTC------AGCCAGTCTGTCACTGCCTAT 1852
                 |||:::      |||          |||          ::::::   |||:::
Db        457 ValSerValLeuAlaGlyProPheAspSerLeuPheGlyGlyGlyAsnLeuProAlaPhe 476

Qy       1853 ATGGTG---------TCTGCCGCAGGCCTGGGTCTGGTCGCCATT 1888
               :::|||          |||||||||     |||::::::||||||
Db        477 ValValGlyAlaIleSerAlaAlaIleSerGlyValLeuAlaIle 491
```

RESULT 31
Q43653
ID Q43653 PRELIMINARY; PRT; 516 AA.
AC Q43653;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Sucrose transport protein.
GN SUT1.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. DESIREE; TISSUE=LEAF;
RX MEDLINE=94146554; PubMed=8312741;
RA Riesmeier J., Willmitzer L., Frommer W.B.;
RT "Potato sucrose transporter expression in minor veins indicates a role
RT in phloem loading.";
RL Plant Cell 5:1591-1598(1993).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; X69165; CAA48915.1; -.
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR TIGRFAMs; TIGR01301; GPH_sucrose; 1.
KW Transmembrane.
SQ SEQUENCE 516 AA; 54831 MW; 4FD06C095E49A377 CRC64;

Alignment Scores:

Pred. No.:	2.18e-13	Length:	516
Score:	333.00	Matches:	133
Percent Similarity:	41.93%	Conservative:	93
Best Local Similarity:	24.68%	Mismatches:	201
Query Match:	5.19%	Indels:	112
DB:	10	Gaps:	21

US-09-759-143-110 (1-3410) x Q43653 (1-516)

```
Qy        278 CCCACTATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTC 337
              |||            :::||||||                              ::::::
Db         24 ProLeuAlaProSerLysLeuTrp----------------------------LysIle 33

Qy        338 TTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTAT 397
              ::::::|||       :::     |||::::::      |||      ::::::
Db         34 IleValValAlaSerIleAlaAlaGlyValGlnPheGlyTrpAlaLeuGlnLeuSerLeu 53

Qy        398 GTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGC 457
              :::  |||  :::     :::|||:::      ||||||   :::   :::
Db         54 LeuThrProTyrValGlnLeuLeuGlyIleProHisLysPheAlaSerPheIleTrpLeu 73

Qy        458 ATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGG 517
                 ||||||:::   |||::::::      |||:::::::|||   ||||||:::
Db         74 CysGlyProIleSerGlyMetIleValGlnProValValGlyTyrTyrSerAspAsnCys 93

Qy        518 CGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTG 577
                 |||:::||||||||||||||||||   |||   :::      ::::::::
Db         94 SerSerArgPheGlyArgArgArgProPheIleAlaAlaGlyAlaAlaLeuValMetIle 113

Qy        578 AGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTG----------------- 619
              :::::|||||||||         |||:::   |||   |||
Db        114 AlaValPheLeuIle---------GlyPheAlaAlaAspLeuGlyHisAlaSerGlyAsp 130

Qy        620 ---CTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTG 676
                 |||      :::|||      |||:::   ::::::|||       :::
Db        131 ThrLeuGlyLysGlyPheLysProArgAlaIleAlaValPheValValGlyPheTrpIle 150

Qy        677 CTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTC 736
              ||||||            :::         |||   |||||||||:::||||||
Db        151 LeuAspValAlaAsnAsnMetLeuGlnGlyProCysArgAlaLeuLeuAlaAspLeuSer 170

Qy        737 CGGGACCCGGAC---CACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTT 793
                                |||   |||   :::   ::::::|||   :::::::::
Db        171 GlyGlyLysSerGlyArgMetArgThrAlaAsnAlaPhePheSerPhePheMetAlaVal 190

Qy        794 GGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGG-------------------- 832
              |||      |||||||||      :::
Db        191 GlyAsnIleLeuGlyTyrAlaAlaGlySerTyrSerHisLeuPheLysValPheProPhe 210

Qy        833 ---GACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT---GGC 886
                    |||   |||         |||            :::   |||   |||
Db        211 SerLysThrLysAlaCysAspMetTyrCysAlaAsnLeuLysSerCysPhePheIleAla 230

Qy        887 CTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG--- 943
              :::      |||:::   ||||||   :::|||   ||||||:::            |||
Db        231 IlePheLeuLeuLeuSerLeuThrThrIleAlaLeuThrLeuValArgGluAsnGluLeu 250

Qy        944 GCAGCGCTGGGCCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
                                         |||       |||   |||:::
Db        251 ProGluLysAspGluGlnGluIleAspGluLysLeuAlaGlyAlaGly----------- 266

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
                           ::::::::::::    |||       |||   :::       |||
Db        267 -----------LysSerLysValProPhe------PheGlyGluIlePheGlyAlaLeu 280

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
                 :::|||            ||||||   :::      |||   :::
Db        281 LysGluLeu------------ProArgProMetTrpIleLeuLeuLeuValThrCysLeu 296

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              :::|||:::|||         |||   |||:::   ||||||:::::::   :::   :::
Db        297 AsnTrpIleAlaTrpPheProPhePheLeuTyrAspThrAspTrpMetAlaLysGluVal 316

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              :::   |||            :::   |||         |||   ||||||   |||
Db        317 PheGlyGly------------GlnValGlyAsp------AlaArgLeuTyrAspLeuGly 330

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||   |||::::::||||||   ||||||      :::            ||||||
Db        331 ValArgAlaGlyAlaMetGlyLeuLeuLeuGlnSerValValLeuGlyPheMetSerLeu 350

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
```

```
Db       1241 LysGlyGluThrGlyAspValGlyGlnMetGlyProProGlyProProGlyProArgGly 1260

Qy        314 GCAGGCGGCTCACCCACAGCCTCTGGACCATAGTGGGCCA-------------------- 275
                 |||:::|||  |||  ||||||  ||||||
Db       1261 ProSerGlyAlaProGlyAlaAspGlyPro-GlnGlyProProGlyGlyIleGlyAsnPr 1280

Qy        274 ----------GGCGGGTAGGGCTCAGGGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCT 225
                        |||   |||   |||  :::|||  |||  |||
Db       1280 oGlyAlaValGlyGluLysGlyGluProGlyGluAlaGlyGluProGlyLeuProGlyGl 1300

Qy        224 CGGCTCT-----------------------------------------GCTCCAGAAGC 207
                |||                                          |||
Db       1300 uGlyGlyProLeuGlyProLysGlyGluArgGlyGluLysGlyGluValGlyProSerGl 1320

Qy        206 TGCGGCCTCTCCTCCTTGCTGCCGC--------------CAACTGCCTAGGAATCAGCC 162
                ||||||  ||||||       :::              |||:::      ||
Db       1320 yAlaAlaGlyProProGlyProLysGlyProProGlyAspAspGlyProLysGlySerPr 1340

Qy        161 AGGCGCCCATTTCTGCCAGCCCTTTGGT-----------GCCGGTCCAGCTTCTCAGCC 114
                ||||             |||  |||           ||||||      :::||
Db       1340 oGly--------------ProValGlyPheProGlyAspProGlyProProGlyGluPr 1355

Qy        113 CATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTCATCACTCAGATC 54
                |           ||||||             ||||||            :::|
Db       1355 o------Gly-ProAlaGlyGlnAspGlyProProGlyAspLysGlyAspAspGlyGluP 1373

Qy         53 CTGGCCGAGGCGCGCGGCTGTCACCCGGA 25
                |||||:::            |||||||||
Db       1373 roGlyGlnThrGly------SerProGly 1380
```

```
RESULT 30
Q9ZTB9
ID   Q9ZTB9      PRELIMINARY;      PRT;   512 AA.
AC   Q9ZTB9;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Sucrose transporter.
GN   SUT1.
OS   Apium graveolens (Celery).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids II; Apiales; Apiaceae; Apium.
OX   NCBI_TaxID=4045;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LEAF;
RA   Noiraud N., Delrot S., Lemoine R.;
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR   EMBL; AF063400; AAC99332.1; -.
DR   InterPro; IPR003662; sub_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   TIGRFAMs; TIGR01301; GPH_sucrose; 1.
KW   Transmembrane.
SQ   SEQUENCE   512 AA;  54520 MW;  4D76A7854A7DF903 CRC64;

Alignment Scores:
Pred. No.:              2.17e-13       Length:          512
Score:                  333.00         Matches:         139
Percent Similarity:     40.54%         Conservative:    86
Best Local Similarity:  25.05%         Mismatches:      198
Query Match:            5.19%          Indels:          132
DB:                     10             Gaps:            21

US-09-759-143-110 (1-3410) x Q9ZTB9 (1-512)

Qy        332 CAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATC 391
                :::|||:::||||||        :::       |||::::::        |||     :::
Db         33 LysLeuIleLeuValAlaAlaIleAlaAlaGlyValGlnPheGlyTrpAlaLeuGlnLeu 52

Qy        392 ACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTG 451
                :::    :::    |||    :::      :::|||:::      |||:::       :::
Db         53 SerLeuLeuThrProTyrValGlnLeuLeuGlyIleProHisLysTrpAlaAlaTyrIle 72

Qy        452 CTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGAC 511
                    ||||||:::   |||::::::      |||::::::|||      ||||||
Db         73 TrpLeuCysGlyProIleSerGlyMetLeuValGlnProIleValGlyTyrTyrSerAsp 92

Qy        512 CACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATC 571
                   :::      :::||||||||||||||||||||    :::    :::       :::
Db         93 ArgCysGlnSerSerPheGlyArgArgArgProPheIleAlaSerGlyAlaGlyCysVal 112

Qy        572 CTGCTGAGCCTCTTTCTCATCCCAAGGGCC---------GGCTGGCTAGCAGGGCTGCTG 622
                :::|||:::   ||||||       |||     |||:::  ||||||
Db        113 AlaIleSerValIleLeuIleGlyPheAlaAlaAspIleGlyTyrLysAlaGly------ 130

Qy        623 TGCCCGGATCCC--------------AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGC 667
                     |||               :::|||    :::    :::   ::::::|||
Db        131 ------AspAspMetAsnLysThrLeuLysProArgAlaValThrValPheValIleGly 148

Qy        668 GTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCT 727
                   :::||||||          :::             |||     |||||||||:::
Db        149 PheTrpIleLeuAspValAlaAsnAsnMetLeuGlnGlyProCysArgAlaLeuLeuAla 168

Qy        728 GACCTCTTCCGG---GACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATG 784
                ||||||         |||          |||   |||   ::::::|||   |||
Db        169 AspLeuCysAsnGlyAspThrArgArgMetArgSerAlaAsnAlaLeuTyrArgPhePhe 188

Qy        785 ATCAGTCTTGGGGGCTGCCTGGGC-------------------------TACCTCCTG 817
                :::::::::|||      ||||||                          |||   |||
Db        189 MetAlaValGlyAsnIleLeuGlyAsnAlaAlaGlySerTyrAsnAsnLeuTyrLysLeu 208

Qy        818 CCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGC 877
                                |||   |||           |||             :::   |||
Db        209 PheProPheSer---LysThrHisAlaCysAspLeuTyrCysAlaAsnLeuLysSerCys 227

Qy        878 ---CTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 934
                   :::       ::::::   |||||||||:::|||   :::|||   |||      |||
Db        228 PheIleIleSerIleValLeuLeuIlePheIleThrValLeuAlaLeuThr------Val 245

Qy        935 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 994
                   |||              |||   |||       |||             :::
Db        246 ValArgGluLysGlnTrpSerProAspGluAlaAspGluGluProProSerSerGlyLys 265

Qy        995 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1054
                   |||                                          |||   ||||||
Db        266 IlePro--------------------------------ValPheGlyGluLeuLeu 273

Qy       1055 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1114
                   |||       |||               ||||||   :::      |||
Db        274 ArgAlaLeuLysAspLeu-----------ProArgProMetLeuMetLeuLeuAlaVal 289

Qy       1115 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 1174
                      :::|||:::|||            |||   ||||||   ||||||::::::|||
Db        290 ThrCysLeuAsnTrpIleAlaTrpPheProPheIleLeuPheAspThrAspTrpMetGly 309

Qy       1175 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 1234
                   :::|||   |||            |||         |||            :::   |||
Db        310 ArgGluIleTyrGlyGlyThr------AlaGlyGlnGly-----------LysLeuTyr 323

Qy       1235 GATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTC 1294
                |||:::||||||||||:::||||||||||||||||||      |||      :::   :::
Db        324 AspGlnGlyValArgValGlySerLeuGlyLeuLeuLeuAsnSerValValLeuGlyLeu 343

Qy       1295 TTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCC 1354
                   |||:::   ::::::   ||||||:::         |||         :::   |||
Db        344 ThrSerIleAlaValGluTyrLeuValArgGlyValGlyGlyValLysIle---LeuTrp 362

Qy       1355 AGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTG 1414
                   :::      |||   :::|||   |||                     :::   |||
Db        363 GlyLeuValAsnPheLeuLeuAlaIleGlyLeuVal--------------MetThrVal 377

Qy       1415 GTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTAC 1474
                |||   :::      |||
Db        378 ValValSerLysValAla---------------------------------------- 383
```

```
Qy       2159 AGGCAGCCCTAGAGACTGGGGAGAGAGG-----------------------AGAGGGAC 2124
Db        657 -GlyLeuProGlyAspAspGlyGluArgGlyAspAspGlyGluValGlyProArgGlyLe 676
Qy       2123 GCCCCAGCCCCCAGCTGTGCAG------CTACGCACCTCAGCAGCACAGGGTGGCAGCAG 2070
Db        676 uProGlyGluProGlyProArgGlyLeuLeuGlyProLysGlyProProGlyProProGl 696
Qy       2069 AGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGC 2010
Db        696 yProPro---------Gly-ValThrGlyMetAspGlyGlnProGly---ProLysGlyA 712
Qy       2009 TAACAGGA---GCGGGGAGCTGGGACCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGG 1953
Db        712 snValGlyProGlnGlyGluProGlyProProGlyGlnGlnGlyAsnProGly------- 729
Qy       1952 AAGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAA 1893
Db        730 ----------AlaGlnGlyLeu-ProGlyProGlnGlyAlaIleGlyProProGlyGlu 745
Qy       1892 AGTAAATGGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGC 1833
Db        746 LysGlyProLeuGlyLysProGlyLeuProGlyMetProGlyAlaAspGlyProProGly 765
Qy       1832 TGAGCTGGACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGG---ACAGCAGGAAGG 1776
Db        766 HisProGlyLys-----------------GluGlyProProGlyGluLysGlyGlyGln 779
Qy       1775 CACTATCCAGGATGGCGAGGTCCAGGCAGATGCCCCGGCCCGGAA-------------- 1731
Db        780 GlyProProGlyProGlnGlyProIleGlyTyrProGlyProArgGlyValLysGlyAla 799
Qy       1731 ------------------------------------------------------------ 1731
Db        800 AspGlyIleArgGlyLeuLysGlyThrLysGlyGluLysGlyGluAspGlyPheProGly 819
Qy       1730 ------------------------------------------CCACCCTGGCCTCGG 1716
Db        820 PheLysGlyAspMetGlyIleLysGlyAspArgGlyGluIleGlyProProGlyProArg 839
Qy       1715 TGGGCTCACCCACCACCACACGTACGGAGACATCACAGGCAGAGGCCCCGCAGAGCGCGG 1656
Db        840 --GlyGlu----------------------AspGlyProGluGlyProLysGlyArgG 851
Qy       1655 GTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGGGAAG---GGAG 1599
Db        851 lyGlyProAsnGlyAspProGlyProLeuGlyProThrGlyGluLysGlyLysLeuGlyV 871
Qy       1598 CTCCAGGCTTAGGG------CCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCTAG 1545
Db        871 alProGlyLeuProGlyTyrProGlyArgGln-------------------------G 882
Qy       1544 CACCTCCAGTGTCCCCTCGGTATT------TGGGCAGGAACACCTGCTTCTCCCGGTGGT 1491
Db        882 lyProLysGlySer-IleGlyPheProGlyPheProGlyAlaAsnGlyGluLysGlyGly 901
Qy       1490 AGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGG 1431
Db        902 ArgGlyThrPro-----------GlyLysProGly----------------------- 909
Qy       1430 CGGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACCGGCAGCCACAG 1371
Db        910 ----------------------------------------------------ProArg 911
Qy       1370 GGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCTGCACCAGCCGGT 1311
Db        912 GlyGln-------------------------------------ArgGlyProThrGly 918
Qy       1310 CCATGACCAGAGAGA--------------------AGACCAGGG---AGATGGCGCACT 1275
Db        919 ProArgGlyGluArgGlyProArgGlyIleThrGlyLysProGlyProLysGlyAsnSer 938
Qy       1274 GCAGGAACAGCCCCA---GGCTGCCCATCCGAACGCCTTCATCATAGTGTCTCCGGGCCT 1218
Db        939 GlyGlyAspGlyProAlaGlyProProGlyGluArg-----------------GlyPro 952
Qy       1217 CGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTACAGCCCCTCGCCCACGAAATCCGTGT 1158
Db        953 AsnGlyProGlnGlyProThrGlyPheProGlyProLysGlyProProGlyProProGly 972
Qy       1157 AAAACAGCGTGAAGGTCATGAGTGCCA------------------------------ 1131
Db        973 LysAspGly-----------LeuProGlyHisProGlyGlnArgGlyGluThrGlyPhe 988
Qy       1130 TCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACA 1071
Db        989 GlnGlyLysThrGlyProProGlyProProGlyVal---ValGlyProGlnGlyProThr 1007
Qy       1070 ------GCTGGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGG--- 1020
Db       1008 GlyGluThrGlyProMetGlyGluArgGlyHisProGlyProProGlyProProGlyGlu 1027
Qy       1019 ---------CCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTTCTG 969
Db       1028 GlnGlyLeuProGlyValAlaGlyLysGluGlyThrLysGlyAspPro----------- 1043
Qy        968 CTGGCTCGGTGGGGCCCAGCG---CTGCCTCCTCAG------CCACCAGCAGTGTGGCTG 918
Db       1044 -----------GlyProAlaGlyLeuProGlyLysAspGlyProPro----------- 1055
Qy        917 CTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGCACT-------------- 873
Db       1056 -----------GlyLeuArgGlyPheProGlyAspArgGlyLeuProGlyProValGly 1071
Qy        872 -------------------------CCTCCTGGGTGCCCAGGT------AGGGGGCCA 846
Db       1072 AlaLeuGlyLeuLysGlySerGluGlyProProGlyProProGlyProAlaGlySerPro 1091
Qy        845 GGGCACTGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGACTGA 786
Db       1092 GlyGluArgGlyProAla------GlyAlaAlaGlyProIleGlyIlePro--------- 1106
Qy        785 TCATGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGTCCGGGTCCCGGAAGAGGTCAG 726
Db       1107 -----------------GlyArgProGlyProGlnGlyPro----------------- 1114
Qy        725 AGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGC 666
Db       1115 ---ProGlyProAlaGlyGluLysGlyValProGlyGluLysGlyProGlnGlyPro-Al 1133
Qy        665 CCAGGATGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTAGCC 606
Db       1133 aGlyArgAsp------GlyLeuGlnGlyPro---ValGlyLeuProGlyProAlaGlyPr 1150
Qy        605 AGCCGGCCCTTGGGATGAG----------------------AAAGAGGCTCAGCAGGA 570
Db       1150 oValGlyProProGlyGluAspGlyAspLysGlyGluIleGlyGluProGlyGlnLysGl 1170
Qy        569 TGCCCAAGGACAGTGCCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGT 510
Db       1170 ySerLysGlyAspLysGlyGluGlnGlyProProGly-ProThrGlyPro---GlnGlyP 1189
Qy        509 CACTGGCTGAGCCT---AGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCA 453
Db       1189 roIleGlyGlnProGlyProSerGlyAlaAspGlyGluPro------GlyProArgGlyG 1207
Qy        452 GCACCATGGTCATGAACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCACATAGG 393
Db       1207 lnGlnGlyLeu------PheGlyGlnLysGlyAspGluGlySerArgGly---------- 1221
Qy        392 TGATGCCTGCGGCCAAACACACCTCCAGGC---CAAAGGTTAGCAGGTTGACCAGCAAGA 336
Db       1222 --PheProGly-ProProGlyProValGlyLeuGlnGlyLeuProGlyProProGlyGlu 1240
Qy        335 GCTGGGCTTTCCGGT------------------------------------GCCGCA 315
```

```
Qy        338 ---AGAGCTGGGCTTTCCGGTGCCGCAGCAGGCGGCTCACCCACAGCCTCTGGACCATAG 282
              |||  ||||||  |||              |||  |||      :::||||||
Db       1134 GlyArgAspGlyLeuGlnGlyProVal------GlyLeuProGlyProAlaGlyPro-Va 1151

Qy        281 TGGG---CCAGGCGGGTAGGGCTCAGGGGGCCGTTCAGGCACTCCA---GAACTGCTTCG 228
              |||    ||||||       |||      |||       |||   |||    :::
Db       1151 lGlyProProGlyGluAspGlyAspLysGlyGluIleGlyGluProGlyGlnLysGlySe 1171

Qy        227 TCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAACTGCCTAGGAA 168
                 |||        |||              |||         |||   |||
Db       1171 rLysGlyAspLysGlyGluGlnGlyProProGlyProThrGlyProGlnGlyProThrGl 1191

Qy        167 TCAGCCAGGCGCCCATTTCTGCC------------------------------------- 145
                |||||||| |||      |||
Db       1191 yGlnProGly-ProSerGlyAlaAspGlyGluProGlyProArgGlyGlnGlnGlyLeuP 1211

Qy        144 -------------------------------------AGCCCTTTGGTGCCGGTCCAGC 123
                                                   |||      ||||||   |
Db       1211 heGlyGlnLysGlyAspGluGlySerArgGlyPheProGlyProProGlyProValGlyL 1231

Qy        122 TTCTCAGCCCATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTCATC 63
              ||                      |||      |||      :::||||||
Db       1231 euGlnGlyLeuProGlyProProGlyGluLysGlyGlu---ThrGlyAspValGlyGlnM 1250

Qy         62 ACTCAGATCCTGGC-----------CGAGGCGCGCGGCTGTCACCCGGAGCC 22
                    ||||||          ||||||          :::|||||||||
Db       1250 etGlyProProGlyProProGlyProArgGlyProSerGlyAlaProGlyAla 1267
```

```
RESULT 29
Q60467
ID   Q60467      PRELIMINARY;      PRT;  1840 AA.
AC   Q60467;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Pro-alpha-1 type V collagen.
OS   Cricetulus longicaudatus (Long-tailed hamster) (Chinese hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Cricetulus.
OX   NCBI_TaxID=10030;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92105142; PubMed=1722213;
RA   Greenspan D.S., Cheng W., Hoffman G.G.;
RT   "The pro-alpha-1(V) collagen chain: Complete primary structure,
RT   distribution of expression, and comparison with the pro-alpha-1(XI)
RT   collagen chain.";
RL   J. Biol. Chem. 266:24727-24733(1991).
DR   EMBL; M76730; AAA37002.1; -.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR000885; Fib_collagen_C.
DR   InterPro; IPR001791; Laminin_G.
DR   InterPro; IPR001230; Prenyl_site.
DR   InterPro; IPR003129; TSPN.
DR   Pfam; PF01410; COLFI; 1.
DR   Pfam; PF01391; Collagen; 18.
DR   Pfam; PF02210; TSPN; 1.
DR   ProDom; PD000007; Collagen; 2.
DR   ProDom; PD002078; Fib_collagen_C; 1.
DR   SMART; SM00038; COLFI; 1.
DR   SMART; SM00282; LamG; 1.
DR   SMART; SM00210; TSPN; 1.
DR   PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
KW   Collagen.
SQ   SEQUENCE   1840 AA;  184174 MW;  32C56821EF64CE75 CRC64;

Alignment Scores:
Pred. No.:              2.15e-13      Length:         1840
Score:                  334.50        Matches:        302
Percent Similarity:     29.66%        Conservative:   59
Best Local Similarity:  24.82%        Mismatches:     392
Query Match:            5.37%         Indels:         464
DB:                     11            Gaps:           68

US-09-759-143-110 (1-3410) x Q60467 (1-1840)
```

```
Qy       3068 GGTCCTGCAGTAGCTCCAAACAGGGTTGTGGAGCTGGTGGGGAAAGTTGGGGGTAGGGGA 3009
              ||||||   :::       |||:::                |||     |||  ||||||
Db        433 GlyProGlyMetProAlaAsnGlnAspThrIleTyrGluGlyIleGlyGlyProArgGly 452

Qy       3008 AAGTTGGGG---GTAGGGGAAATTTTGGGCAGTGCCTTCATCAGCCCAGTCCTAGAGAGA 2952
              :::   |||      ||||||            |||   |||   |||    :::
Db        453 GluLysGlyGlnLysGlyGlu---------ProAlaIleIleGluProGlyMet---Leu 468

Qy       2951 GTAGAGGGG-----------------------------------------------AGT 2940
              :::||||||                                               :::
Db        469 IleGluGlyProProGlyProGluGlyProAlaGlyLeuProGlyProProGlyThrThr 488

Qy       2939 GGAAGTGGGGGGAACCAGGCTGGGCCAAGAGAAGAGGGGTGGTTAGGGAAGCCGTTGAGA 2880
              |||       |||            |||   |||   |||       |||:::|||
Db        489 GlyProThrGlyGlnValGlyAspProGlyGluArgGlyProProGlyArgProGlyLeu 508

Qy       2879 CCTGAAGCC---------------CCACCCTCTACCTTCCTTCAACACCCTAACCTTGGG 2835
              |||   |||                ||||||   |||    |||       |||
Db        509 ProGlyAlaAspGlyLeuProGlyProProGlyThrMetLeuMetLeuPro--------- 525

Qy       2834 TAACAGCATTTGGAATTATCATTTGGGATGAGTAGAATTTCCAAGGTCCTGGGTTAGGCA 2775
                                   ||||||                          |||
Db        526 ---------------PheArgPheGly-----------------------GlyGlyGly 532

Qy       2774 TTTTGGGGGGCCAGA--------------CCCCAGGAGAAGAAGATTCTGGCAATGATC 2730
                 |||:::::                ||||||    :::     |||:::::
Db        533 AspAlaGlySerLysGlyProMetValSerAlaGlnGluSerGlnAlaGlnAlaIleLeu 552

Qy       2729 AGCCCAATGACCAGCTATCTCAGGGGA----------------------------CCT 2700
                                ||||||||                            |||
Db        553 GlnGlnAlaArgLeuAlaLeuArgGlyProAlaGlyProMetGlyLeuThrGlyArgPro 572

Qy       2699 GATTGTTGGGGATCCCCCACCCTACCCAAATATTAGACACCAACACAGAAAAGCTAGCAA 2640
                   ||| ||||||         ::::::   :::
Db        573 GlyProMetGly-ProProGlySerGlyGlyLeuLysGlyGluProGlyAspMetGlyPr 592

Qy       2639 TGGATTCCCTTCTACTTTGTTAAATAAATAAGTTAAATATTTAAATGCCTGTGTCTCTGT 2580
                    |||                                                   ||
Db        592 oGlnGlyProArg------------------------------------------GlyVa 598

Qy       2579 GATGGCAACAGAAGGACCAACAGGCCACATCCTGATAAAAGGTAAGAGGGGGGTGGATCA 2520
              |          ||||||||||||        :::||||:::   |||  |||
Db        598 lGlnGlyProProGlyProThrGlyLysProGlyArgArgGlyArgAlaGlySerAsp-- 617

Qy       2519 GCAAAAAGACAGTGCTGTGGGCTGAGGGGACCTGGTTCTTGTGTGTTGCCCCTCAGGACT 2460

Db        617 ------------------------------------------------------------ 617

Qy       2459 CTTCCCCTACAAATAAGTCATATGTTCAAATCCCATGGAGGAGTGTTTCATCCTAGAAAC 2400

Db        617 ------------------------------------------------------------ 617

Qy       2399 TCCCATGCAAGAGCTACATTAAACGAAGCTGCAGGTTAAGGGGCTTAGAGATGGGAAACC 2340
                                                    |||   |||   |||   ||
Db        618 ------------------------------------GlyAlaArg---GlyMetPr 623

Qy       2339 AGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCTCAACTAGGAGG 2280
              ||||           :::   ||||||          |||         |||   ||
Db        623 oGlyGln---------ThrGlyProLysGly---AspArgGlyPheAspGlyLeuAlaGl 639

Qy       2279 CTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCAGAGTCCCCGCATTCCAGTGCATG 2220
              |                                  |||   |||     |||
Db        639 y--------------------------LeuProGlyGluLysGlyHisArgGlyAspPr 650

Qy       2219 GAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGCCTCCAGTC 2160
                ||||||||   |||
Db        650 oGlyProSerGlyProPro--------------------------------------- 656
```

```
                  ||||              :::  ||||||          |||          |||  ||
Db      623 oGlyGln---------ThrGlyProLysGly---AspArgGlyPheAspGlyLeuAlaGl 639

Qy     2279 CTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCAGAGTCCCCGCATTCCAGTGCATG 2220
            |                                  |||  |||   |||
Db      639 y--------------------------LeuProGlyGluLysGlyHisArgGlyAspPr 650

Qy     2219 GAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGCCTCCAGTC 2160
                 |||||||||   |||
Db      650 oGlyProSerGlyProPro------------------------------------------ 656

Qy     2159 AGGCAGCCCTAGAGACTGGGGAGAGAGG------------------------AGAGGGAC 2124
             |||  |||  |||  |||||||||                        ||||||
Db      657 -GlyLeuProGlyAspAspGlyGluArgGlyAspAspGlyGluValGlyProArgGlyLe 676

Qy     2123 GCCCCAGCCCCCAGCTGTGCAG------CTACGCACCTCAGCAGCACAGGGTGGCAGCAG 2070
             |||    |||       :::     |||                 |||
Db      676 uProGlyGluProGlyProArgGlyLeuLeuGlyProLysGlyProProGlyProProGl 696

Qy     2069 AGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGC 2010
               |||          |||  |||      ||||||||||||  |||  |||
Db      696 yProPro---------Gly-ValThrGlyMetAspGlyGlnProGly---ProLysGlyA 712

Qy     2009 TAACAGGAGCGGGGAGCTGGGACCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAG 1950
                |||                |||   |||   |||    |||
Db      712 snValGly---------------ProGln-GlyGluProGlyProPro------------ 722

Qy     1949 TTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGT 1890

Db      722 ------------------------------------------------------------ 722

Qy     1889 AAATGGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGA 1830
                  :::   ::::||||||   :::       |||       |||        |||
Db      723 GlyGlnGlnGlyAsnProGlyAlaGlnGlyLeuProGlyProGlnGlyAlaIleGlyPro 742

Qy     1829 GCTGGACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCACTAT 1770
               |||:::        |||        :::|||  ||||||     |||
Db      743 ProGlyGluLysGlyPro---------LeuGlyLysProGlyLeuProGly------Met 757

Qy     1769 CCAGGATGGCGAGGTCCAGGCAGATGCCCCGGCCCGGAACCACCCTGGCCTCGGTGGGCT 1710
            ||||||        ||||||              |||    |||||||||      |||
Db      758 ProGlyAlaAspGlyPro---------------Pro-GlyHisProGlyLysGluGly-- 771

Qy     1709 CACCCACCACCACACGTACGGAGACATCACAGGCAGAGGCCCCGCAGAGCGCGGGTGGAG 1650
                                                     |||   |||:::  ||||||
Db      772 --------------------------------------ProProGlyGluLys-GlyGlyG 779

Qy     1649 GTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGGG---------------- 1606
             |||              ||||||   |||      |||:::|||
Db      779 lnGly------------ProProGlyProGlnGlyProIleGlyTyrProGlyProArgG 795

Qy     1605 -----AAGGGAGCTCCAGGCTTA---------------GGCCTGGCAGGAAGCTGGTCA 1566
                 |||||||||  |||:::                      |||   |||
Db      795 lyValLysGlyAlaAspGlyIleArgGlyLeuLysGlyThrLysGlyGluLys-GlyGlu 814

Qy     1565 TCAGGCTGTCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGTATTTGGGCA------GGA 1512
              |||   |||                       :::      :::||||||       |||
Db      815 AspGlyPhePro---------GlyPheLysGlyAspMetGlyIleLysGlyAspArgGly 831

Qy     1511 ACACCTGCTTCTCCCGGTGGTAGAGGGAGG---------------CCAGTGTGTAGGGCA 1467
                         ||||||  ||||||              |||       |||
Db      832 GluIleGlyProProGlyProArgGlyGluAspGlyProGluGlyProLysGlyArgGly 851

Qy     1466 GGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCA 1407
             |||       |||                                    |||   |||
Db      852 GlyProAsnGly------------------------------------AspProGlyPro 859

Qy     1406 CACTGTGGGACAGGCATGTGGCACCGGCAGCCACAGGGAAAGCTGCCACACTGGCCAAAT 1347
                        ||||||                     :::|||||||||
Db      860 LeuGlyProThrGly---------------GluLysGlyLysLeu--------------- 869

Qy     1346 AGACTGCTCGAGTGCCGAATCGCTGCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGG 1287
                                     |||        |||               ||||||
Db      870 ---------------------GlyValProGlyLeuPro---------GlyTyrProGly 879

Qy     1286 AGATGGCGCACTGCAGGAACAGCC------CCAGGCTGCCCATCCGAACGCCTTCATCAT 1233
            |||              |||:::            ||||||   |||
Db      880 ArgGlnGlyProLysGlySerIleGlyPheProGlyPheProGlyAlaAsnGlyGluLys 899

Qy     1232 AGTGTCTCCGGG---CCTCGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTACAGCCCCT 1176
                       |||   |||       |||                |||     |||   |||
Db      900 GlyGlyArgGlyThrProGlyLysProGlyProArgGlyGlnArgGlyProThrGlyPro 919

Qy     1175 CGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGCTGCACAGCT 1116
            |||                                                |||      |||
Db      920 ArgGlyGluArgGly-------------------------ProArgGlyIleThrGly 930

Qy     1115 CAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACAGCTGGTGCAGCCGGG 1056
            :::|||          |||                              :::|||    |||
Db      931 LysProGlyProLysGly----------------------AsnSerGlyGlyAspGly 942

Qy     1055 GAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGGCCCGGCATGGACAGCAGTGGGGCG 996
               ||||||  ||||||        |||:::||||||
Db      943 ProAlaGlyProProGlyGluArgGlyProAsnGlyPro--------------------- 955

Qy      995 ACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCTCCTCAG 936
                  :::|||||||||                    ||||||    ||||||
Db      956 ------GlnGlyProThrGlyPhePro------------GlyProLysGlyProProGly 969

Qy      935 CCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGC 876
            ||||||                       :::    |||       ||||||||||||
Db      970 ProProGly--------------LysAspGlyLeuProGlyHisProGlyGlnArgGly 984

Qy      875 ACT-----------------------CCTCCTGGGTGCCCAGGT------------AGG 852
                                      |||||||||  ||||||               :::
Db      985 GluThrGlyPheGlnGlyLysThrGlyProProGlyProProGlyValValGlyProGln 1004

Qy      851 GGGCCA----------------------------GGGCACTGGTGTCCCAGTCAATGG 822
            ||||||                            ||||||
Db     1005 GlyProThrGlyGluThrGlyProMetGlyGluArgGlyHis------------------ 1018

Qy      821 CAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCATAGACAGAGTAGG 762
                    |||   ||||||   |||    :::                     |||
Db     1019 ------ProGlyProProGlyProProGlyGlu-----------------GlnGlyLeu 1030

Qy      761 CCTGGCGACAGTGGT----------------------CCGGGTCCC---GGAAGAGGTCAG 726
            ||||||   :::|||                     |||||||||    |||
Db     1031 ProGlyAlaAlaGlyLysGluGlyThrLysGlyAspProGlyProAlaGlyLeuProGly 1050

Qy      725 AGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCA--- 669
            :::   |||||||||                    ||||||  |||          |||
Db     1051 LysAspGlyProProGlyLeuArgGlyPheProGlyAspArgGlyLeuProGlyProVal 1070

Qy      668 ----------------------------CGCCCAGGATGAGCAGTGCCAGCTCCAGGG 639
                                          ||||||         |||
Db     1071 GlyAlaLeuGlyLeuLysGlySerGluGlyProProGly---------ProProGlyPro 1087

Qy      638 GCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGC 579
            |||     ||||||                                    |||||||||
Db     1088 AlaGlySerProGly-------------------------------------GluArgGly 1095

Qy      578 TCAGCAGGATGCCCAAGGACAGTGCCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCAC 519
               ||||||
Db     1096 ProAlaGly--------------------------------------------------- 1098

Qy      518 GCCAGTGGTCACTGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAA 459
            |||:::|||               :::|||       ||||||||||||
Db     1099 AlaAlaGly---------ProIleGlyIleProGlyArgProGlyPro------------ 1111

Qy      458 TGCCCAGCACCATGGTCATGAACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCA 399
                                          ||||||   ||||||   |||
Db     1112 -------------------------GlnGlyProProGlyProAlaGlyGluLysGly 1122

Qy      398 CATAGGTGATGCCTGCGGCCAAACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCA 339
                                      ||||||:::                  ||||||
Db     1123 --------------------------ValProGlyGluGluGlyProGlnGlyProAla 1133
```

```
Db      1208 eu------PheGlyGlnLysGlyAspGluGlySerArgGly-----------PheProG 1222

Qy       383 CGGCCAAACACACCTCCAGGC---CAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTT 327
                 |||    |||   |||   |||   |||   |||   |||         |||
Db      1222 ly-ProProGlyProValGlyLeuGlnGlyLeuProGlyProProGlyGluLysGlyGlu 1241

Qy       326 TCCGGT--------------------------------------GCCGCAGCAGGCGGC 306
             :::|||                                             |||
Db      1242 ThrGlyAspValGlyGlnMetGlyProProGlyProProGlyProArgGlyProSerGly 1261

Qy       305 TCACCCACAGCCTCTGGACCATAGTGGGCCA---------------------------- 275
             :::|||   |||   ||||||   ||||||
Db      1262 AlaProGlyAlaAspGlyPro-GlnGlyProProGlyGlyIleGlyAsnProGlyAlaVa 1281

Qy       274 -GGCGGGTAGGGCTCAGGGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCTCGGC----- 221
              |||     |||      |||   :::|||   |||   |||         |||
Db      1281 lGlyGluLysGlyGluProGlyGluAlaGlyAspProGlyLeuProGlyGluGlyGlyPr 1301

Qy       220 -----------------------------------TCTGCTCCAGAAGCTGCGGCCTC 198
                                                 :::   |||     ||||||
Db      1301 oLeuGlyProLysGlyGluArgGlyGluLysGlyGluAlaGlyProSerGlyAlaAlaGl 1321

Qy       197 TCCTCCTTGCTGCCGC---------------CAACTGCCTAGGAATCAGCCAGGCGCCCA 153
              ||||||     :::                      |||:::      ||||||
Db      1321 yProProGlyProLysGlyProProGlyAspAspGlyProLysGlySerProGly----- 1339

Qy       152 TTTCTGCCAGCCCTTTGGT-----------GCCGGTCCAGCTTCTCAGCCCATGCTCAA 105
                          |||   |||               ||||||     :::|||
Db      1340 ---------ProValGlyPheProGlyAspProGlyProProGlyGluPro------Gl 1354

Qy       104 CACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTCATCACTCAGATCCTGGCCGAG 45
                 ||||||                 ||||||                :::||||||:::
Db      1354 y-ProAlaGlyGlnAspGlyProProGlyAspLysGlyAspAspGlyGluProGlyGlnT 1374

Qy        44 GCGCGCGGCTGTCACCCGGA 25
                         |||||||||
Db      1374 hrGly------SerProGly 1378
```

RESULT 28
Q9JI03

```
ID   Q9JI03      PRELIMINARY;      PRT;  1840 AA.
AC   Q9JI03;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Alpha 1 type V collagen.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE DAWLEY;
RX   MEDLINE=20428740; PubMed=10852920;
RA   Chernousov M.A., Rothblum K., Tyler W.A., Stahl R.C., Carey D.J.;
RT   "Schwann cells synthesize type V collagen that contains a novel alpha
RT   4 chain. molecular cloning, biochemical characterization, and high
RT   affinity heparin binding of alpha4(V) collagen.";
RL   J. Biol. Chem. 275:28208-28215(2000).
DR   EMBL; AF272662; AAF76433.1; -.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR000885; Fib_collagen_C.
DR   InterPro; IPR001791; Laminin_G.
DR   InterPro; IPR001230; Prenyl_site.
DR   InterPro; IPR000508; SigPTase.
DR   InterPro; IPR003129; TSPN.
DR   Pfam; PF01410; COLFI; 1.
DR   Pfam; PF01391; Collagen; 17.
DR   Pfam; PF02210; TSPN; 1.
DR   ProDom; PD000007; Collagen; 2.
DR   ProDom; PD002078; Fib_collagen_C; 1.
DR   SMART; SM00038; COLFI; 1.
DR   SMART; SM00282; LamG; 1.
DR   SMART; SM00210; TSPN; 1.
DR   PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
DR   PROSITE; PS00761; SPASE_I_3; UNKNOWN_1.
KW   Collagen.
SQ   SEQUENCE   1840 AA;  183986 MW;  AD38F5FF886B923C CRC64;

Alignment Scores:
Pred. No.:              1.99e-13      Length:         1840
Score:                  335.00        Matches:        290
Percent Similarity:     28.98%        Conservative:   47
Best Local Similarity:  24.94%        Mismatches:     356
Query Match:            5.38%         Indels:         470
DB:                     11            Gaps:           64

US-09-759-143-110 (1-3410) x Q9JI03 (1-1840)

Qy      3068 GGTCCTGCAGTAGCTCCAAACAGGGTTGTGGAGCTGGTGGGGAAAGTTGGGGGTAGGGGA 3009
             ||||||   :::      |||:::                  |||   |||   ||||||
Db       433 GlyProGlyMetProAlaAsnGlnAspThrIleTyrGluGlyIleGlyGlyProArgGly 452

Qy      3008 AAGTTGGGG---GTAGGGGAAATTTTGGGCAGTGCCTTCATCAGCCCAGTCCTAGAGAGA 2952
             :::   |||      ||||||           |||   |||   |||   :::
Db       453 GluLysGlyGlnLysGlyGlu---------ProAlaIleIleGluProGlyMet---Leu 468

Qy      2951 GTAGAGGGG-------------------------------------------------AGT 2940
             :::||||||                                                  :::
Db       469 IleGluGlyProProGlyProGluGlyProAlaGlyLeuProGlyProProGlyThrThr 488

Qy      2939 GGAAGTGGGGGGAACCAGGCTGGGCCAAGAGAAGAGGGGTGGTTAGGGAAGCCGTTGAGA 2880
             |||      |||               |||   |||   |||      |||:::|||
Db       489 GlyProThrGlyGlnMetGlyAspProGlyGluArgGlyProProGlyArgProGlyLeu 508

Qy      2879 CCTGAAGCC---------------CCACCCTCTACCTTCCTTCAACACCCTAACCTTGGG 2835
             |||   |||                  ||||||   |||   |||      |||
Db       509 ProGlyAlaAspGlyLeuProGlyProProGlyThrMetLeuMetLeuPro--------- 525

Qy      2834 TAACAGCATTTGGAATTATCATTTGGGATGAGTAGAATTTCCAAGGTCCTGGGTTAGGCA 2775
                                   ||||||                         |||
Db       526 ---------------PheArgPheGly-----------------------GlyGlyGly 532

Qy      2774 TTTTGGGGGGCCAGA---------------CCCCAGGAGAAGAAGATTCTGGCAATGATC 2730
                |||::::::                  ||||||   :::      |||::::::
Db       533 AspAlaGlySerLysGlyProMetValSerAlaGlnGluSerGlnAlaGlnAlaIleLeu 552

Qy      2729 AGCCCAATGACCAGCTATCTCAGGGGA-------------------------------CCT 2700
                               |||||||||                            |||
Db       553 GlnGlnAlaArgLeuAlaLeuArgGlyProAlaGlyProMetGlyLeuThrGlyArgPro 572

Qy      2699 GATTGTTGGGGATCCCCCACCCTACCCAAATATTAGACACCAACACAGAAAAGCTAGCAA 2640
                      ||| ||||||               ::::::   :::
Db       573 GlyProMetGly-ProProGlySerGlyGlyLeuLysGlyGluProGlyAspMetGlyPr 592

Qy      2639 TGGATTCCCTTCTACTTTGTTAAATAAATAAGTTAAATATTTAAATGCCTGTGTCTCTGT 2580
                      |||                                               ||
Db       592 oGlnGlyProArg-------------------------------------------GlyVa 598

Qy      2579 GATGGCAACAGAAGGACCAACAGGCCACATCCTGATAAAAGGTAAGAGGGGGGTGGATCA 2520
             |               |||||||||||            :::|||:::  |||   |||
Db       598 lGlnGlyProProGlyProThrGlyLysProGlyArgArgGlyArgAlaGlySerAsp-- 617

Qy      2519 GCAAAAAGACAGTGCTGTGGGCTGAGGGGACCTGGTTCTTGTGTGTTGCCCCTCAGGACT 2460

Db       617 ------------------------------------------------------------ 617

Qy      2459 CTTCCCCTACAAATAAGTCATATGTTCAAATCCCATGGAGGAGTGTTTCATCCTAGAAAC 2400

Db       617 ------------------------------------------------------------ 617

Qy      2399 TCCCATGCAAGAGCTACATTAAACGAAGCTGCAGGTTAAGGGGCTTAGAGATGGGAAACC 2340
                                                     |||   |||   |||   ||
Db       618 ---------------------------------------GlyAlaArg---GlyMetPr 623

Qy      2339 AGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCTCAACTAGGAGG 2280
```

```
Qy    2279 CTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCAGAGTCCCCGCATTCCAGTGCATG 2220
Db     637 y---------------------------LeuProGlyGluLysGlyHisArgGlyAspPr 648
Qy    2219 GAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGCCTCCAGTC 2160
Db     648 oGlyProSerGlyProPro----------------------------------------- 654
Qy    2159 AGGCAGCCCTAGAGACTGGGGAGAGAGG------------------------AGAGGGAC 2124
Db     655 -GlyIleProGlyAspAspGlyGluArgGlyAspAspGlyGluValGlyProArgGlyLe 674
Qy    2123 GCCCCAGCCCCCAGCTGTGCAG------CTACGCACCTCAGCAGCACAGGTGGCAGCAG 2070
Db     674 uProGlyGluProGlyProArgGlyLeuLeuGlyProLysGlyProProGlyProProGl 694
Qy    2069 AGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGC 2010
Db     694 yProPro---------Gly-ValThrGlyMetAspGlyGlnProGly---ProLysGlyA 710
Qy    2009 TAACAGGA---GCGGGGAGCTGGGACCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGG 1953
Db     710 snValGlyProGlnGlyGluProGlyProProGlyGlnGlnGlyAsnProGly------- 727
Qy    1952 AAGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAA 1893
Db     728 -----------AlaGlnGlyLeu-ProGlyProGlnGlyAlaIleGlyProProGlyGlu 743
Qy    1892 AGTAAATGGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGC 1833
Db     744 LysGlyProLeuGlyLysProGlyLeuProGlyMetProGlyAlaAspGlyProProGly 763
Qy    1832 TGAGCTGGACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGG---ACAGCAGGAAGG 1776
Db     764 HisProGlyLys------------------GluGlyProProGlyGluLysGlyGlyGln 777
Qy    1775 CACTATCCAGGATGGCGAGGTCCAGGCAGATGCCCCGGCCCGGAA--------------- 1731
Db     778 GlyProProGlyProGlnGlyProIleGlyTyrProGlyProArgGlyValLysGlyAla 797
Qy    1731 ------------------------------------------------------------ 1731
Db     798 AspGlyIleArgGlyLeuLysGlyThrLysGlyGluLysGlyGluAspGlyPheProGly 817
Qy    1730 ---------------------------------------------CCACCCTGGCCTCGG 1716
Db     818 PheLysGlyAspMetGlyIleLysGlyAspArgGlyGluIleGlyProProGlyProArg 837
Qy    1715 TGGGCTCACCCACCACCACACGTACGGAGACATCACAGGCAGAGGCCCCGCAGAGCGCGG 1656
Db     838 --GlyGlu------------------------AspGlyProGluGlyProLysGlyArgG 849
Qy    1655 GTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGGGAAG---GGAG 1599
Db     849 lyGlyProAsnGlyAspProGlyProLeuGlyProThrGlyGluLysGlyLysLeuGlyV 869
Qy    1598 CTCCAGGCTTAGGG------CCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCTAG 1545
Db     869 alProGlyLeuProGlyTyrProGlyArgGln---------------------------G 880
Qy    1544 CACCTCCAGTGTCCCCTCGGTATT------TGGGCAGGAACACCTGCTTCTCCCGGTGGT 1491
Db     880 lyProLysGlySer-IleGlyPheProGlyPheProGlyAlaAsnGlyGluLysGlyGly 899
Qy    1490 AGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGG 1431
Db     900 ArgGlyThrPro------------GlyLysProGly------------------------ 907
Qy    1430 CGGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACCGGCAGCCACAG 1371
Db     908 ------------------------------------------------------ProArg 909
Qy    1370 GGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCTGCACCAGCCGGT 1311

Db     910 GlyGln---------------------------------------ArgGlyProThrGly 916
Qy    1310 CCATGACCAGAGAGA---------------------AGACCAGGG---AGATGGCGCACT 1275
Db     917 ProArgGlyGluArgGlyProArgGlyIleThrGlyLysProGlyProLysGlyAsnSer 936
Qy    1274 GCAGGAACAGCCCCA---GGCTGCCCATCCGAACGCCTTCATCATAGTGTCTCCGGGCCT 1218
Db     937 GlyGlyAspGlyProAlaGlyProProGlyGluArg------------------GlyPro 950
Qy    1217 CGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTACAGCCCCTCGCCCACGAAATCCGTGT 1158
Db     951 AsnGlyProGlnGlyProThrGlyPheProGlyProLysGlyProProGlyProProGly 970
Qy    1157 AAAACAGCGTGAAGGTCATGAGTGCCATCCAGCTGCACAGCTCAGCCACGAAGAGCCGGC 1098
Db     971 Lys------------------------AspGlyLeuProGlyHisPro------------ 978
Qy    1097 GCAGGGTGCGGGGCATGCGGCAGCACAGCT------GGTGCAGCCGGGGAAGCAGGGCGC 1044
Db     979 ---------GlyGlnArgGlyGluThrGlyPheGlnGlyLysThrGlyProProGlyPro 995
Qy    1043 CCAGGTTCCGGAAAGCCAAGCGGCCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGG 984
Db     996 ProGlyValValGlyProGlnGlyPro------------ThrGlyGluThr------Gly 1009
Qy     983 CCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCTCCTCAGCCACCAGCAGTG 924
Db    1010 Pro------------MetGlyGluArgGlyHisProGlyProProGlyProProGly--- 1024
Qy     923 TGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGCACT--------- 873
Db    1025 ------------GluGlnGlyLeuProGlyAlaAlaGlyLysGluGlyThrLysGlyAsp 1040
Qy     872 ------CCTCCTGGGTGCCCAGGTAGG---GGGCCA------------------------ 846
Db    1041 ProGlyProAlaGlyLeuProGlyLysAspGlyProProGlyLeuArgGlyPheProGly 1060
Qy     845 GGGCACTGGTGTCCCAGTCAATGGCAGGCA------------------GGAGGTAGCCCA 804
Db    1061 AspArgGlyLeuProGlyProValGlyAlaLeuGlyLeuLysGlySerGluGlyProPro 1080
Qy     803 GGCAGCCCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGTCCG 744
Db    1081 GlyProProGlyPro------------------AlaGlySerProGlyGluArgGlyPro 1094
Qy     743 GGT------------------------CCCGGAAGAGGTCAGAGAGCAGGGCCTCCA--- 711
Db    1095 AlaGlySerAlaGlyProIleGlyIleProGlyArgProGlyProGlnGlyProProGly 1114
Qy     710 ------GTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGCCCAGGATGA 657
Db    1115 ProAlaGlyGluLysGlyLeuProGlyGluLysGlyProGlnGlyPro-AlaGlyArgAs 1134
Qy     656 GCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGGCCC 597
Db    1134 p------GlyLeuGlnGlyPro---ValGlyLeuProGlyProAlaGlyProValGlyPr 1151
Qy     596 TTGGGATGAG------------------------AAAGAGGCTCAGCAGGATGCCCAAGG 561
Db    1151 oProGlyGluAspGlyAspLysGlyGluIleGlyGluProGlyGlnLysGlySerLysGl 1171
Qy     560 ACAGTGCCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTG 501
Db    1171 yAspLysGlyGluGlnGlyProProGly-ProThrGlyPro---GlnGlyProIleGlyG 1190
Qy     500 AGCCT---AGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCATGG 444
Db    1190 lnProGlyProSerGlyAlaAspGlyGluPro------GlyProArgGlyGlnGlnGlyL 1208
Qy     443 TCATGAACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTG 384
```

```
Db       900 alMetValGlyProProGlyAlaLysGlyGluLysGlyAlaProGlyAspLeuAlaGlyA 920

Qy       308 --------GGCTCACCCACAGCCTCTGGACCATAGTGGG-------------------- 278
                     |||    |||    |||    |||       |||
Db       920 spLeuLeuGlyGluProGlyAlaLysGly-AspArgGlyLeuProGlyProArgGlyGlu 939

Qy       277 --------------------------CCAGGCGGGTAGGGCTCAGGGGGCCGTTCAGGC 245
                                       ||||||       |||       |||:::   |||
Db       940 LysGlyGluAlaGlyArgAlaGlyGluProGlyAspProGlyGluAspGlyGlnLysGly 959

Qy       244 ACTCCAGAACTGCTTCGTCTCGGC--------------------------TCTGCTCCA 212
               |||         |||   |||                                 |||
Db       960 AlaProGly------ArgLysGlyLeuLysGlyGluProGlyIleGlyValGlnGlyPro 977

Qy       211 GAAGCTGCGGCCTCTCCTCCTTGCTGCCGC---------------------------- 182
                          ||||||       :::
Db       978 ProGlyProThrGlyProProGlyMetLysGlyAspValGlySerProGlyAlaProGly 997

Qy       181 --------------------CAACTGCCTAGG---------AATCAGCCAGGCGCCCAT 152
                                 ||||||           ||||||||| |||::
Db       998 ValValGlyPheProGlyGlnThrGlyProArgGlyGluThrGlyGlnProGly-ProVa 1017

Qy       151 TTCTGCCAGCCCTTTGGTG----------------------------CCGGTCCAGCT 122
             :             |||                              |||
Db      1017 lGlyGluArgGlyLeuAlaGlyProProGlyArgGluGlyAlaProGlyPro-------- 1034

Qy       121 TCTCAGCCCATGCTCAACACCTGCTGCTGTGGGG------CACCTCAGTGGGGACACGTC 68
               |||   |||          |||   :::||||||           ||||||
Db      1035 -LeuGlyProProGlyProProGlySerValGlyAlaProGlyAlaSerGlyLeuLysGl 1054

Qy        67 TCATCACTCAGATCCTGGC-----------------CGAGGCGCGCGGCTGTCACCCGG 26
                         |||||||||                  ||||||   |||      |||||
Db      1054 yAspLysGlyAspProGlyThrGlyLeuProGlyProArgGlyGluArgGlyGluProGl 1074

Qy        25 A 25
             |
Db      1074 y 1074
```

RESULT 27
O88207

```
ID   O88207     PRELIMINARY;      PRT;  1838 AA.
AC   O88207;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Collagen a1(V).
GN   COL5A1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98250615; PubMed=9582436;
RA   Wu Y.L., Sumiyoshi H., Khaleduzzaman M., Ninomiya Y., Yoshioka H.;
RT   "cDNA sequence and expression of the mouse alpha1(V) collagen gene.";
RL   Biochim. Biophys. Acta 1397:275-284(1998).
DR   EMBL; AB009993; BAA28786.1; -.
DR   MGD; MGI:88457; Col5a1.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR000885; Fib_collagen_C.
DR   InterPro; IPR001791; Laminin_G.
DR   InterPro; IPR001230; Prenyl_site.
DR   InterPro; IPR003129; TSPN.
DR   Pfam; PF01410; COLFI; 1.
DR   Pfam; PF01391; Collagen; 18.
DR   Pfam; PF02210; TSPN; 1.
DR   ProDom; PD000007; Collagen; 2.
DR   ProDom; PD002078; Fib_collagen_C; 1.
DR   SMART; SM00038; COLFI; 1.
DR   SMART; SM00282; LamG; 1.
DR   SMART; SM00210; TSPN; 1.
DR   PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
KW   Collagen.
SQ   SEQUENCE   1838 AA;  183691 MW;  7A520B23D1851783 CRC64;

Alignment Scores:
Pred. No.:              1.85e-13      Length:         1838
Score:                  335.50        Matches:        301
Percent Similarity:     29.74%        Conservative:   60
Best Local Similarity:  24.79%        Mismatches:     395
Query Match:            5.39%         Indels:         458
DB:                     11            Gaps:           68

US-09-759-143-110 (1-3410) x O88207 (1-1838)

Qy      3068 GGTCCTGCAGTAGCTCCAAACAGGGTTGTGGAGCTGGTGGGGAAAGTTGGGGGTAGGGGA 3009
             ||||||    :::       |||:::                |||     |||   ||||||
Db       431 GlyProGlyMetProAlaAsnGlnAspThrIlePheGluGlyIleGlyGlyProArgGly 450

Qy      3008 AAGTTGGGG---GTAGGGGAAATTTTGGGCAGTGCCTTCATCAGCCCAGTCCTAGAGAGA 2952
             :::   |||      ||||||                |||   |||   |||     :::
Db       451 GluLysGlyGlnLysGlyGlu---------ProAlaIleIleGluProGlyMet---Leu 466

Qy      2951 GTAGAGGGG-------------------------------------------------AGT 2940
             :::||||||                                                  :::
Db       467 IleGluGlyProProGlyProGluGlyProAlaGlyLeuProGlyProProGlyThrThr 486

Qy      2939 GGAAGTGGGGGGAACCAGGCTGGGCCAAGAGAAGAGGGGTGGTTAGGGAAGCCGTTGAGA 2880
             |||     |||                |||    |||    |||        |||:::|||
Db       487 GlyProThrGlyGlnMetGlyAspProGlyGluArgGlyProProGlyArgProGlyLeu 506

Qy      2879 CCTGAAGCC---------------CCACCCTCTACCTTCCTTCAACACCCTAACCTTGGG 2835
             |||   |||                ||||||    |||    |||         |||
Db       507 ProGlyAlaAspGlyLeuProGlyProProGlyThrMetLeuMetLeuPro--------- 523

Qy      2834 TAACAGCATTTGGAATTATCATTTGGGATGAGTAGAATTTCCAAGGTCCTGGGTTAGGCA 2775
                                     ||||||                          |||
Db       524 ---------------PheArgPheGly-----------------------GlyGlyGly 530

Qy      2774 TTTTGGGGGGCCAGA---------------CCCCAGGAGAAGAAGATTCTGGCAATGATC 2730
                      |||::::::                  ||||||    :::      |||::::::
Db       531 AspAlaGlySerLysGlyProMetValSerAlaGlnGluSerGlnAlaGlnAlaIleLeu 550

Qy      2729 AGCCCAATGACCAGCTATCTCAGGGGA---------------------------CCT 2700
                                ||||||||                            |||
Db       551 GlnGlnAlaArgLeuAlaLeuArgGlyProAlaGlyProMetGlyLeuThrGlyArgPro 570

Qy      2699 GATTGTTGGGGATCCCCCACCCTACCCAAATATTAGACACCAACACAGAAAAGCTAGCAA 2640
                      ||| ||||||                   ::::::    :::
Db       571 GlyProMetGly-ProProGlySerGlyGlyLeuLysGlyGluProGlyAspMetGlyPr 590

Qy      2639 TGGATTCCCTTCTACTTTGTTAAATAAATAAGTTAAATATTTAAATGCCTGTGTCTCTGT 2580
                      |||                                                   ||
Db       590 oGlnGlyProArg-------------------------------------------GlyVa 596

Qy      2579 GATGGCAACAGAAGGACCAACAGGCCACATCCTGATAAAAGGTAAGAGGGGGGTGGATCA 2520
             |                   |||||||||||          :::|||:::    |||    |||
Db       596 lGlnGlyProProGlyProThrGlyLysProGlyArgArgGlyArgAlaGlySerAsp-- 615

Qy      2519 GCAAAAAGACAGTGCTGTGGGCTGAGGGGACCTGGTTCTTGTGTGTTGCCCCTCAGGACT 2460

Db       615 ------------------------------------------------------------ 615

Qy      2459 CTTCCCCTACAAATAAGTCATATGTTCAAATCCCATGGAGGAGTGTTTCATCCTAGAAAC 2400

Db       615 ------------------------------------------------------------ 615

Qy      2399 TCCCATGCAAGAGCTACATTAAACGAAGCTGCAGGTTAAGGGGCTTAGAGATGGGAAACC 2340
                                                      |||    |||    |||    ||
Db       616 ---------------------------------------GlyAlaArg---GlyMetPr 621

Qy      2339 AGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCTCAACTAGGAGG 2280
             ||||     .     :::    ||||||             |||            |||   ||
Db       621 oGlyGln---------ThrGlyProLysGly---AspArgGlyPheAspGlyLeuAlaGl 637
```

```
Qy   2119 -----CAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGAGC 2065
               :::  ||||||                             :::|||          :::|
Db    269 lyValArgGlyProAla-----------------GlyGluLysGlyAspGlnGlyAspP 283

Qy   2064 CACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAACA 2005
          ||         |||   |||   :::      |||:::    |||   |||
Db    283 roGlyGluAspGlyArgAsnGlySerProGlyProSer---GlyProLysGlyAspArg- 301

Qy   2004 GGAGCGGGGAGCTGGGACCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTTC 1945
               ||||||    ||||||          |||||||||   :::     |||:::    :
Db    302 -----GlyGluProGlyPro------ProGlyProProGlyArgLeuValGlyAlaGlyI 318

Qy   1944 TACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAATG 1885
          ::                                               |||:::     |
Db    318 leGly------------------------------------------SerArgAspLysG 324

Qy   1884 GCGACCA------------------------------GACCCA--------------- 1872
          ||:::                                ||||||
Db    324 lyGluProGlyGlnGluGlyProArgGlyProLysGlyAspProGlyProProGlyAlaS 344

Qy   1871 -------------------GGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGCT 1832
                                |||||||||      |||      |||          |||
Db    344 erGlyGluArgGlyIleGluGlyLeuArgGlyProProGlyProGlnGlyAspProGlyV 364

Qy   1831 GAGCTGGACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCACT 1772
                |||                            |||||||||||     ||||||:::
Db    364 alArgGlyProAlaGlyAspLysGlyAspArgGlyProProGlyLeuAspGlyArgAsn- 383

Qy   1771 ATCCAGGATGGCGAGGTCCAGGCAGATGCCCCGGCCCGGAACCACCCTGGCCTCGGTGGG 1712
                |||     |||          |||||||||
Db    384 -----GlyValAspGlyLysProGlyAlaProGlyPro--------------------- 394

Qy   1711 CTCACCCACCACCACACGTACGGAGACATCACAGGCAGAGGCCCCGCAGAGCGCGGGTGG 1652
               |||   ||||||       :::     :::|||       |||          |||
Db    395 -----ProGlyProHis--GlyAlaSerGlyLysAlaGlyAspProGlyArgAspGlyLe 412

Qy   1651 AGGTGGGAGCAGG-----------CCACTG---CCTCCAGCACCCACGTGTCCATTAGG 1607
            |||   |||              |||      ||||||    |||           ||
Db    412 uProGlyLeuArgGlyGluHisGlyProAlaGlyProProGlyProProGlyValProGl 432

Qy   1606 GAAG-----------GGAGCTCCAGGCTTAGGG------------------------- 1585
          ||||            |||   |||||||||
Db    432 yLysThrGlyGluAspGlyLysProGlyLeuAsnGlyLysAsnGlyGluProGlyAspPr 452

Qy   1584 ----------------CCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCTAGCACC 1541
                             |||   |||          ||||||   |||
Db    452 oGlyGluAspGlyArgLysGlyGluLys-GlyAspSerGlyAlaProGlyArgGluGlyP 472

Qy   1540 TCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGTAGAGGGAGGC 1481
                |||   |||         :    |||   |||      ||||||   ||||||    |
Db    472 roAspGlyProLysGly---GluArgGlyAlaProGlyAspProGlyLeuArgGlyProP 491

Qy   1480 CA-----------GTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAG 1433
          ||                  ||||||    .  ||||||    |||   ||||||
Db    491 roGlyLeuProGlyGln-ValGlyProProGlyGlnGlyPheProGlyValProGlyAsn 510

Qy   1432 GGCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACA--------------------- 1395
             |||           |||   |||   |||:::
Db    511 ThrGlyProLysGlyAspArgGlyGluThr--GlySerLysGlyGluGlnGlyLeuProG 530

Qy   1394 --------GGCATGTGGCACCGGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGAC 1343
                  |||:::      |||            |||    ||||||:::       ||||
Db    530 lyGluArgGlyLeu------ArgGlyGluProGlySerLeuProAsn---AlaGluArgL 547

Qy   1342 TGCTCGAG----------TGCCGAATCGCTGCA--------------------------- 1320
          ||||||||                  :::::::::|||
Db    547 euLeuGluThrAlaGlyIleLysValSerAlaLeuArgAspIleValGluThrTrpGlyG 567

Qy   1319 --CCAGCCGGT--------CCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAG 1271
                :::|||          |||   ||||||||||||||||||               |
Db    567 luSerSerGlySerPheLeuProValProGluArgArgProGlyProLysGlyAspProG 587

Qy   1270 GAACA-----------------------GCCCCAGGCTGCCCATCCGAACGCCTTCATC 1235
          ||                             :::   |||   |||    ||||||
Db    587 lyGluArgGlyProProGlyLysGluGlySerIleGlyPheProGlyGluArg------- 604

Qy   1234 ATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTACAGCCCCTC 1175
                 :::   |||   |||                        ||||||   |||
Db    605 --GlyLeuLysGlyAspArgGlyAspProGlyProGlnGlyProProGlyLeuAlaLeuG 624

Qy   1174 GCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGCTGCACAGCTC 1115
                |||   |||                              |||            :
Db    624 lyGluArgGlyProProGly--------------------ProProGlyLeuAlaGlyG 637

Qy   1114 AGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACAGCTGGTGCAGCCGGGG 1055
          ::|||   :::      |||         |||       |||         ||||||:::||||||
Db    637 luProGlyLysProGlyIlePro---GlyLeuProGly---ArgAlaGlySerAlaGlyG 655

Qy   1054 AAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGGCCCGGCATGGACAGCAGTGGGGCGA 995
          |||||||||||||||                     |||            |||
Db    655 luAlaGlyArgProGlyGluArgGlyGluArgGlyGluLysGlyAspArgGlyGluGlnG 675

Qy    994 CAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCTCCTCAGC 935
            |||   |||          |||                ||||||     ||||||    |
Db    675 lyArgAspGlyLeuProGlyLeuPro-----------GlyProProGlyProProGlyP 691

Qy    934 CACCAGCAGTG-----------------TGGCTGCTACGCAGGTGAGGAAGATGAGGGT 893
          ||                            |||   |||      |||         |||
Db    691 roLysValAlaIleAspGluGlnGlyProGlyLeuSerArgGluGlnGlyProProGlyL 711

Qy    892 GAGCAGGCCAAAGAGGCACTCCT-------------------------------CCTG 866
             |||    :::|||   |||                                     |
Db    711 euLysGlyAlaLysGlyGluProGlySerAspGlyAspHisGlyProLysGlyAspLysG 731

Qy    865 GGTGCCCAGGTAGGGGGCCAGGGCACTGGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCC 806
           ||    ||||||                     |||   :::              ||||||
Db    731 lyAlaProGlyIleLysGlyAspGlnGlyGluProGlyLysArgGlyHisAspGlySerP 751

Qy    805 CAGGCAGCCCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGT- 747
           |||||     |||      :::                  :::     :::||||||       |||
Db    751 roGlyLeuProGlyGluArgGlyValAlaGlyProGluGlyLysProGlyLeuGlnGlyP 771

Qy    746 ----------CCGGGTCCCGGAAGAGGTCAG---AGAGCAGGGCCTCCAGTGGAGTGAA 701
                    |||||||||        |||           |||||||||
Db    771 roArgGlyThrProGlyProAlaGlyGlyHisGlyAspProGlyProPro--------G 788

Qy    700 GCACACCTGGCCACAGAAGTCCAGCAGCCCCACGC------------------------ 666
               ||||||          ||||||     |||:::
Db    788 lyAlaProGlyLeuAlaGlyProAlaGlyProGlnGlyProSerGlyLeuLysGlyGluP 808

Qy    665 -----------------CCAGGATGAGCAGTGCCAGCTCCA---GGGGCC--------- 636
                            ||||||        :::|||    |||    ||||||
Db    808 roGlyGluThrGlyProProGlyArgGlyLeuProGlyProThrGlyAlaValGlyLeuP 828

Qy    635 --TGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCT 578
                  ||||||      :::      |||:::                        :::||||
Db    828 roGlyProProGlyProSerGlyLeuValGly-----------------ProGlnGlyS 842

Qy    577 CAGCAGGATGCCCAAGGACAGTG--------------------CCCAGATGAAGGGCCG 539
          ||   |||   |||      |||                     |||        |||   |
Db    842 erProGlyLeuProGlyGlnValGlyGluThrGlyLysProGlyProProGlyArgAspG 862

Qy    538 GCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTAGGAGCGGGACACAGAC 479
          ||                                ||||||                      |||   |||
Db    862 ly------------------ThrSerGlyLys---------------AspGlyGluArgG 871

Qy    478 CAGGCCCAGCACTGGACCAATGCCCAGCACCATGGTCATGAACTTCTCCTCTACCCCCAC 419
            ||||||    :::              |||                          |||   |||
Db    871 lyGlyProGlyValProGlyLeuProGly---------------LeuProGlyProValG 886

Qy    418 TTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCCAAACACACCTCCAGGCCAAA 359
            |||      |||                                |||       |||||||||
Db    886 lyProLysGlyGluProGly------------------ProValGlyAlaProGlyGlnV 900

Qy    358 GGTTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGTGCC---------GCAGCAGGC- 309
            :::   |||   |||         |||         ||||||                ||||||
```

```
                 |||     :::     |||   :::          |||   ||||||
Db      364 ValLysArgLeuTrpGlyIleValAsnPheLeuLeu---------AlaIleCysLeuGly 380

Qy     1400 CACAGTGTGGCCGTGGTG---------------------------------------- 1417
               :::|||   |||
Db      381 LeuThrValLeuValThrLysLeuAlaGlnHisSerArgGlnTyrAlaProGlyThrGly 400

Qy     1418 ----------------------------------ACAGCTTCAGCCGCCCTCACCGGGTTC 1444
                                                   :::   |||||||||   |||
Db      401 AlaLeuGlyAspProLeuProProSerGluGlyIleLysAlaGlyAlaLeuThrLeuPhe 420

Qy     1445 ACCTTCTCAGCCCTGCAGATC---------------CTGCCCTACACACTGGCCTCCCTC 1489
            :::         :::   :::                :::|||:::   |||||||||:::
Db      421 SerValLeuGlyValProLeuAlaIleThrTyrSerIleProPheAlaLeuAlaSerIle 440

Qy     1490 TACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGC 1549
            :::                                           |||   ||||||
Db      441 Phe----------------------------------SerSerThrSerGlyAlaGly 448

Qy     1550 AGTGAGGACAGCCTG-----------------------------------ATGACCAGC 1573
                     ||||||                                       |||
Db      449 GlnGlyLeuSerLeuGlyValLeuAsnLeuAlaIleValIleProGlnMetPheValSer 468

Qy     1574 TTCCTGCCAGGCCCTAAGCCTGGA-------------------GCTCCCTTCCCTAAT 1612
                 |||   ||||||                               |||
Db      469 ValLeuSerGlyProTrpAspAlaLeuPheGlyGlyGlyAsnLeuProAlaPheValVal 488

Qy     1613 GGACACGTGGGTGCTGGAGGCAGTGGC--------------CTGCTCCCACCTCCACCC 1657
             |||   |||   |||      ||||||              |||||||||   ||||||
Db      489 GlyAlaValAlaAlaLeuAlaSerGlyIleLeuSerIleIleLeuLeuProSerProPro 508
```

RESULT 26
Q60444

```
ID   Q60444      PRELIMINARY;      PRT;  1549 AA.
AC   Q60444;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Type VII collagen (Fragment).
OS   Cricetulus griseus (Chinese hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Cricetulus.
OX   NCBI_TaxID=10029;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LUNG;
RX   MEDLINE=93271985; PubMed=8499916;
RA   Greenspan D.S.;
RT   "The carboxyl-terminal half of type VII collagen, including the non-
RT   collagenous NC-2 domain and intron/exon organization of the
RT   corresponding region of the COL7A1 gene.";
RL   Hum. Mol. Genet. 2:273-278(1993).
DR   EMBL; L06863; AAA36968.1; -.
DR   HSSP; P00981; 1DTK.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   Pfam; PF01391; Collagen; 22.
DR   Pfam; PF00014; Kunitz_BPTI; 1.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000007; Collagen; 9.
DR   ProDom; PD000222; Kunitz_BPTI; 1.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 1.
KW   Collagen; Serine protease inhibitor.
FT   NON_TER       1      1
SQ   SEQUENCE   1549 AA;  148117 MW;  5D646391E0C29292 CRC64;

Alignment Scores:
Pred. No.:              1.8e-13       Length:          1549
Score:                  335.50        Matches:         324
Percent Similarity:     30.87%        Conservative:    61
Best Local Similarity:  25.98%        Mismatches:      419
Query Match:            5.39%         Indels:          444
DB:                     11            Gaps:            73

US-09-759-143-110 (1-3410) x Q60444 (1-1549)

Qy     3039 GGAGCTGGTGGGGAAAGTTGGGGGTAGGGGAAAGTTGGGGGTAGGGGAAATTTTGGGCAG 2980
            |||:::   |||                  |||:::   |||   :::|||      |||
Db       36 GlySerProGlyProGlnGlyProAlaGlyArgAlaGlyGluLysGlyGluLysGlyAsp 55

Qy     2979 TGCCTTCAT--------------------CAGCCC----------------------- 2965
            |||                           ||||||
Db       56 CysGluAspGlyAlaProGlyLeuProGlyGlnProGlyAlaProGlyGluProGlyLeu 75

Qy     2964 --------------------AGTCCTAGAGAGAGTAGAGGGGAGTGGAAGTGGGGGAA 2926
                                   |||:::      ||||||:::            |||
Db       76 ArgGlyThrProGlyIleThrGlyProLysGlyAspArgGlyGln------ThrGlyThr 93

Qy     2925 CCAGGCTGGGCCAAGAGAAGAGGGGTGGTTAGGGAAGCCGTTGAGACCTGAAGCCCCACC 2866
            ||||||   |||   |||:::|||      |||   |||           :::   |||
Db       94 ProGly-GluProGlyGluLysGlyGluArgGlyProPro-GlyProValGlyProGlnG 113

Qy     2865 CTCTACCTTCCTTCAACACCCTAACCTTGGGTAACAGCATTTGGAATTATCATTTGGGAT 2806
               ||||||                   |||                   |||   |||
Db      113 lyLeuProGly------------ValAlaGly------------------HisProGly- 122

Qy     2805 GAGTAGAATTTCCAAGGTCCTGGGTTAGGCATTTTGGGGGGCCAGACCCCAGGAGAAGAA 2746
               ||||||   |||                      |||   |||   |||      ||||
Db      123 --ValGluGlyProGlu------------------GlyProProGlyProAlaGlyArgA 136

Qy     2745 GATTCTGGCAATGATCAGCCCAATGACCAGCTATCTCAGGGGACCTGATTGTTGGGGATC 2686
            ||     :::               |||                           ||||||
Db      136 rgGlyGluLysGlyGluProGlyArgPro-----------------------GlyAspP 148

Qy     2685 CCCCACCCTACCCA---AATATTAGACACCAACACAGAAAAGCTAGCAATGGATTCCCTT 2629
            ||       |||               :::       |||          |||   |||
Db      148 roAlaValGlyProGlyGlyAlaGlyAlaLysGlyGluLysGlyAspAlaGlyLeuPro- 167

Qy     2628 CTACTTTGTTAAATAAATAAGTTAAATATTTAAATGCCTGTGTCTCTGTGATGGCAACAG 2569

Db      167 ------------------------------------------------------------ 167

Qy     2568 AAGGACCAACAGGCCACATCCTGATAAAAGGTAAGAGGGGGGTGGATCAGCAAAAAGACA 2509
              ||||||   |||          |||||||||::::::|||
Db      168 --GlyProArgGlyAlaAlaGlyIleLysGlyGluGlnGly------------------ 180

Qy     2508 GTGCTGTGGGCTGAGGGGACCTGGTTCTTGTGTGTTGCCCCTCAGGACTCTTCCCCTACA 2449
                                                    |||   ||||||
Db      181 --------------------------------------ProProGlyLeu---------- 184

Qy     2448 AATAAGTCATATGTTCAAATCCCATGGAGGAGTGTTTCATCCTAGAAACTCCCATGCAAG 2389
                  :::      :::|||      :::
Db      185 -----AlaLeuProGlyAspProGlyProLys--------------------------- 193

Qy     2388 AGCTACATTAAACGAAGCTGCAGGTTAAGGGGCTTAGAGATGGGAAACCAGGTGACTGAG 2329
                                                       |||   |||||||||
Db      194 ----------------------------------------GlyAspProGlyAsp---- 198

Qy     2328 TTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCTCAACTAGGAGGCTAGCTGTTAA 2269
                          :::   |||   :::|||:::::           |||
Db      199 -------------ArgGlyPro---IleGlyLeuThrGlyArgAlaGly---------P 210

Qy     2268 CCCTGAGCCTGGGTAATCCACCTGCAGAGTCCCCGCATTCCAGTGCATGGAGCCCTTCTG 2209
                                 ||||||   |||                     |||   |
Db      210 roThrGlyAspSerGlyProProGlyGluLysGlyAspProGlyArgProGlyProProG 230

Qy     2208 GCCTCCCTGTATAAGTCCAG-----------------ACTGAAACCCCCTTGGAAGGCC 2167
            ||   ||||||              :::              |||      :::||||||
Db      230 ly---ProValGlySerArgGlyArgAspGlyGluValGlyGluLysGlyValGluGlyA 249

Qy     2166 TC------CAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGG------AGAGGGACGCCC- 2120
                        |||   |||         |||||||||      ||||||   |||
Db      249 snProGlyAspProGlyLeuProGlyLysAlaGlyGluArgGlyLeuArgGlyAlaProG 269
```

```
Qy       1481 GCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGA 1540
                        |||:::   :::
Db        384 ---------GlnHisGlnArgGlnHis--------------------------------- 389

Qy       1541 GGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCT 1600

Db        389 ------------------------------------------------------------ 389

Qy       1601 CCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCG 1660
                                          |||   :::|||   |||   ||||||   |||
Db        390 --------------------------GlyAlaAsnGlyGlnLeuLeuProProSerAla 400

Qy       1661 CTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAG 1720
                    |||       |||      :::|||:::      ::::::|||   |||
Db        401 ------GlyValLysAlaGlyAlaLeuSerLeuPheSerIleLeuGlyIleProLeuSer 418

Qy       1721 GCCAGGGTG---GTTCCG--------------------------------GGCCGGGGC 1744
                           :::|||                               |||   |||
Db        419 IleThrPheSerIleProPheAlaLeuAlaSerIleTyrSerSerGlySerGlyAlaGly 438

Qy       1745 ATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCC 1804
                    |||   |||   :::|||:::   |||   ::::::   |||:::      |||
Db        439 GlnGlyLeuSerLeuGlyValLeuAsnLeuAlaIleValValProGlnMetIleValSer 458

Qy       1805 CTGTTTATGGGCTCCATTGTCCAGCTC------AGCCAGTCTGTCACTGCCTATATGGTG 1858
              :::      |||               |||         ::::::   |||::::::|||
Db        459 ValLeuAlaGlyProPheAspSerLeuPheGlyGlyGlyAsnLeuProAlaPheValVal 478

Qy       1859 ---------TCTGCCGCAGGCCTGGGTCTGGTCGCCATT 1888
                       |||||||||      |||:::::||||||
Db        479 GlyAlaIleSerAlaAlaIleSerGlyValLeuAlaIle 491
```

RESULT 25
O04077
```
ID   O04077      PRELIMINARY;      PRT;   523 AA.
AC   O04077;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Sucrose transport protein.
GN   SUT.
OS   Vicia faba (Broad bean).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Vicieae; Vicia.
OX   NCBI_TaxID=3906;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=COTYLEDON;
RA   Weber H., Borisjuk L., Heim U., Sauer N., Wobus U.;
RT   "A role for sugar transporters during seed development: Molecular
RT   characterization of a hexose and a sucrose carrier in faba bean
RT   seeds.";
RL   Plant Cell 0:0-0(0).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR   EMBL; Z93774; CAB07811.1; -.
DR   InterPro; IPR003662; sub_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   TIGRFAMs; TIGR01301; GPH_sucrose; 1.
KW   Transmembrane.
SQ   SEQUENCE   523 AA;  55229 MW;  F8EBF170212D191A CRC64;
```

Alignment Scores:

Pred. No.:	1.12e-13	Length:	523
Score:	337.50	Matches:	134
Percent Similarity:	40.38%	Conservative:	76
Best Local Similarity:	25.77%	Mismatches:	185
Query Match:	5.26%	Indels:	125
DB:	10	Gaps:	20

US-09-759-143-110 (1-3410) x O04077 (1-523)

```
Qy        332 CAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATC 391
              ::::::::::::|||      :::      |||::::::         |||   :::
Db         36 LysIleMetValValAlaSerIleAlaAlaGlyValGlnPheGlyTrpAlaLeuGlnLeu 55

Qy        392 ACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTG 451
              :::   :::   |||   :::      :::|||:::            :::      :::
Db         56 SerLeuLeuThrProTyrValGlnLeuLeuGlyIleHisHisThrTrpAlaAlaTyrIle 75

Qy        452 CTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGAC 511
                    ||||||:::   |||::::::   |||:::::::|||         ||||||
Db         76 TrpLeuCysGlyProIleSerGlyMetLeuValGlnProIleValGlyTyrHisSerAsp 95

Qy        512 CACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATC 571
                       |||:::|||||||||||||||||||||   |||   |||:::   :::
Db         96 ArgCysThrSerArgPheGlyArgArgArgProPheIleAlaAlaGlySerIleAlaVal 115

Qy        572 CTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTG------------ 619
                 :::::::::|||||||||         |||:::   |||   |||
Db        116 AlaIleAlaValPheLeuIle---------GlyTyrAlaAlaAspLeuGlyHisSerPhe 132

Qy        620 ---------CTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTG 670
                       |||         ||||||      :::   :::   ::::::|||
Db        133 GlyAspSerLeuAspGlnLysValArgProArgAlaIleGlyIlePheValValGlyPhe 152

Qy        671 GGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGAC 730
                 :::||||||         :::         |||      |||||||||   |||
Db        153 TrpIleLeuAspValAlaAsnAsnMetLeuGlnGlyProCysArgAlaLeuLeuGlyAsp 172

Qy        731 CTCTTCCGG---GACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATC 787
              |||         :::         |||   |||   :::   ::::::|||   :::
Db        173 LeuCysAlaGlyAsnGlnArgLysThrArgAsnAlaAsnAlaPhePheSerPhePheMet 192

Qy        788 AGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC------------------ 829
              ::::::|||      |||||||||         |||
Db        193 AlaValGlyAsnValLeuGlyTyrAlaAlaGlyAlaTyrSerLysLeuTyrHisValPhe 212

Qy        830 ------TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
                       |||   |||            |||         :::   |||   |||
Db        213 ProPheThrLysThrLysAlaCysAsnValTyrCysAlaAsnLeuLysSerCysPhePhe 232

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG---GCTGAG 940
                    |||::::::      ||||||   :::|||   :::   |||:::      :::
Db        233 ------LeuSerIleAlaLeuLeuThrValLeuAlaThrSerAlaLeuIleTyrValLys 250

Qy        941 GAGGCAGCGCTGGGCCCCACCGAGCCAGCA------------GAAGGGCTGTCGGCCCCC 988
              |||   ||||||   |||   :::                  :::|||   |||:::
Db        251 GluThrAlaLeuThrProGluLysThrValValThrThrGluAspGlyGlySerSerGly 270

Qy        989 TCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTG------GCTTTCCGGAACCTG 1042
                 :::                  ||||||      :::|||      ||||||:::   |||
Db        271 GlyMet---------------ProCysPheGlyGlnLeuSerGlyAlaPheLysGluLeu 285

Qy       1043 GGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGGACCCTGCGCCGG 1102
                                                        |||   :::
Db        286 -----------------------------------------LysArgProMetTrpIle 291

Qy       1103 CTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACG 1162
              |||   :::         :::|||:::|||         |||   ||||||   |||
Db        292 LeuLeuLeuValThrCysLeuAsnTrpIleAlaTrpPheProPheLeuLeuPheAspThr 311

Qy       1163 GATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCC 1222
              |||::::::|||:::   :::|||                              ||||||
Db        312 AspTrpMetGlyLysGluValTyr--------------------GlyGlyThrValGly 324

Qy       1223 CGGAGACAC---TATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGC 1279
                    |||   ||||||   |||||||||   |||:::||||||||||   |||
Db        325 GluGlyHisAlaTyrAspMetGlyValArgGluGlyAlaLeuGlyLeuMetLeuAsnSer 344

Qy       1280 GCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGA 1339
                 :::         ||||||   :::|||   |||   :::      |||
Db        345 ValValLeuGlyAlaThrSerLeuGlyValAspIleLeuAlaArgGlyValGly---Gly 363

Qy       1340 GCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCC 1399
```

```
Qy       1550 AGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCT 1609
                                        . ||||||           |||          |||
Db        387 -----------------------PheLeuGlyHisAspLeuPro-----------Pro 394

Qy       1610 AATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGG 1669
                :::|||  |||  |||                             ||||||
Db        395 SerGlyValValIleAla-----------------------------AlaLeuIleVal 404

Qy       1670 GCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCC---------ACCGAG 1720
                 |||       :::  :::   :::          |||
Db        405 PheSerIleLeuGlyIleProLeuAlaIleThrTyrSerValProTyrAlaLeuIleSer 424

Qy       1721 GCCAGGGTGGTTCCG---GGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCC 1777
                 |||:::         |||    |||      |||    :::   :::|||:::   |||
Db        425 ThrArgIleGluSerLeuGlyLeuGlyGlnGlyLeuSerMetGlyValLeuAsnLeuAla 444

Qy       1778 TTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTC------ 1831
                 ::::::   ||||||        ||||||      |||      ||||||
Db        445 IleValIleProGlnValIleValSerLeuGlySerGlyProTrpAspGlnLeuPheGly 464

Qy       1832 ---------AGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTC 1882
                             |||   |||:::          |||:::|||       |||||||||
Db        465 GlyGlyAsnSerProSerLeuAlaValAlaAlaValAlaAlaPheAlaSerGlyLeuVal 484

Qy       1883 GCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGC 1921
              ||||||          :::        |||||||||
Db        485 AlaIleLeuAlaIleProArgSerSerAlaAspLysSer 497
```

RESULT 24
Q9S7Z5
ID   Q9S7Z5   PRELIMINARY;   PRT;   512 AA.
AC   Q9S7Z5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Sucrose transporter SUT2B.
GN   SUT2B OR SUT2A.
OS   Apium graveolens (Celery).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids II; Apiales; Apiaceae; Apium.
OX   NCBI_TaxID=4045;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=MATURE LEAF;
RA   Noiraud N., Lemoine R.;
RT   "Sucrose transporters in celery.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR   EMBL; AF167416; AAD45391.1; -.
DR   EMBL; AF167415; AAD45390.1; -.
DR   InterPro; IPR003662; sub_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   TIGRFAMs; TIGR01301; GPH_sucrose; 1.
KW   Transmembrane.
SQ   SEQUENCE   512 AA; 54426 MW; 2637535216FF1ED2 CRC64;

Alignment Scores:

Pred. No.:	1.03e-13	Length:	512
Score:	338.00	Matches:	140
Percent Similarity:	41.05%	Conservative:	87
Best Local Similarity:	25.32%	Mismatches:	198
Query Match:	5.27%	Indels:	128
DB:	10	Gaps:	21

US-09-759-143-110 (1-3410) x Q9S7Z5 (1-512)

```
Qy        332 CAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATC 391
                 :::|||:::||||||       :::      |||::::::          |||       :::
Db         33 LysLeuIleLeuValAlaAlaIleAlaAlaGlyValGlnPheGlyTrpAlaLeuGlnLeu 52

Qy        392 ACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTG 451
                 :::   :::   |||   :::      :::|||:::      |||:::          :::
Db         53 SerLeuLeuThrProTyrValGlnLeuLeuGlyIleProHisLysTrpAlaAlaTyrIle 72

Qy        452 CTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGAC 511
                       ||||||:::   |||::::::     |||::::::|||       ||||||
Db         73 TrpLeuCysGlyProIleSerGlyMetLeuValGlnProIleValGlyTyrTyrSerAsp 92

Qy        512 CACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATC 571
                    :::      :::||||||||||||||||||||     :::    :::       :::
Db         93 ArgCysGlnSerSerPheGlyArgArgArgProPheIleAlaSerGlyAlaGlyCysVal 112

Qy        572 CTGCTGAGCCTCTTTCTCATCCCAAGGGCC---------GGCTGGCTAGCAGGG---CTG 619
                 :::|||:::   ||||||      |||      |||:::   ||||||
Db        113 AlaIleSerValIleLeuIleGlyPheAlaAlaAspIleGlyTyrLysAlaGlyAspAsp 132

Qy        620 CTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTG 679
                 :::         :::|||      :::         ::::::|||       :::|||
Db        133 MetThrLysThrLeuLysProArgAlaValThrGlyPheValIleGlyPheTrpIleLeu 152

Qy        680 GACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGG 739
              |||          :::         |||      |||||||||:::||||||
Db        153 AspValAlaAsnAsnMetLeuGlnGlyProCysArgAlaLeuLeuAlaAspLeuCysAsn 172

Qy        740 ---GACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGG 796
                 |||          |||   |||   :::   |||:::|||    :::::::::|||
Db        173 GlyAspThrArgArgMetArgSerAlaAsnAlaPheTyrSerPhePheMetAlaValGly 192

Qy        797 GGCTGCCTGGGCTAC----------------------------CTCCTGCCTGCCATT 826
                 |||||||||                                        |||   |||
Db        193 AsnIleLeuGlyTyrAlaAlaGlySerTyrAsnAsnLeuTyrLysLeuPheProPheSer 212

Qy        827 GACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGC---CTCTTT 883
                    |||   |||         |||               :::   |||   :::
Db        213 ------LysThrHisAlaCysAspLeuTyrCysAlaAsnLeuLysSerCysPheIleIle 230

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
                 ::::::   |||||||||:::|||   :::|||   |||      |||      |||
Db        231 SerIleValLeuLeuIlePheIleThrValLeuAlaLeuThr------ValValArgGlu 248

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
                          |||   |||      |||         :::         |||
Db        249 LysGlnTrpSerProAspGluAlaAspGluGluProProSerSerGlyLysIlePro--- 267

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTG---GGCGCCCTGCTTCCCCGG 1060
                                      |||      |||   |||||||||
Db        268 -----------------------ValPheGlyGluLeuPheGlyAlaLeuLysAsp--- 278

Qy       1061 CTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTG 1120
                                     :::||||||   :::         |||   |||
Db        279 --------------------LeuProArgProMetLeuMetLeuLeuValValThrCys 291

Qy       1121 TGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGG 1180
                 :::|||:::|||         |||   ||||||   ||||||:::::|||
Db        292 LeuAsnTrpIleAlaTrpPheProPheIleLeuPheAspThrAspTrpMetGlyArgGlu 311

Qy       1181 CTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAA 1240
              :::|||   |||         |||         |||         :::   ||||||:::
Db        312 IleTyrGlyGlyThr------AlaGlyGlnGly-----------LysLeuTyrAspGln 325

Qy       1241 GGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCT 1300
              ||||||||:::||||||||||||||   |||         :::        :::   |||
Db        326 GlyValArgValGlySerLeuGlyLeuLeuLeuAsnSerValValLeuGlyLeuThrSer 345

Qy       1301 CTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTG 1360
              :::   ::::::   ||||||:::         |||         :::   |||      :::
Db        346 IleAlaValGluTyrLeuValArgGlyValGlyGlyValLysIle---LeuTrpGlyLeu 364

Qy       1361 GCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACA 1420
                 |||   :::|||   |||                           :::   ||||||
Db        365 ValAsnPheLeuLeuAlaIleGlyLeuVal--------------MetThrValValVal 379

Qy       1421 GCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTG 1480
              :::   |||
Db        380 SerLysValAla---------------------------------------------- 383
```

```
Qy        153 ATTTCTGCCAGCCCTTTGGTGCCGGTCCAGCTTCTCAGCCCATGCTCAACACCTGCTGCT 94
              ::         |||          |||||||||    :::
Db        888 snAlaGlyProProGlyProProGlyProAlaGlyLysGlu------------------ 901

Qy         93 GTGGG------GCACCTCAGTGGGGACACGTCTCATCACTCAGATCCTGGCCGA------ 46
                |||         |||:::  |||:::|||               ||||||
Db        902 --GlyGlyLysGlyProArg-GlyGluThrGlyProAlaGlyArgProGlyGluValGly 920

Qy         45 ------------------------GGCGCGCGGCTGTCACCCGGAGCC 22
                                      |||    :::    ||||||||||||
Db        921 ProProGlyProProGlyProAlaGlyGluLysGlySerProGlyAla 936

RESULT 23
Q9SP63
ID   Q9SP63      PRELIMINARY;      PRT;   501 AA.
AC   Q9SP63;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Sucrose transporter.
OS   Vitis vinifera (Grape).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Vitaceae;
OC   Vitis.
OX   NCBI_TaxID=29760;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. UGNI BLANC; TISSUE=GRAPE BERRY;
RA   Ageorges A., Issaly N., Picaud S., Delrot S., Romieu C.;
RT   "Identification and functional expression in yeast of a grape berry
RT   sucrose carrier.";
RL   Plant Physiol. Biochem. 38:177-185(2000).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR   EMBL; AF182445; AAD55269.1; -.
DR   InterPro; IPR003662; sub_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   TIGRFAMs; TIGR01301; GPH_sucrose; 1.
KW   Transmembrane.
SQ   SEQUENCE   501 AA;  54075 MW;  4FD19DFAFE539077 CRC64;

Alignment Scores:
Pred. No.:               9.57e-14     Length:         501
Score:                   338.50       Matches:        163
Percent Similarity:      40.29%       Conservative:   84
Best Local Similarity:   26.59%       Mismatches:     207
Query Match:             5.27%        Indels:         159
DB:                      10           Gaps:           25

US-09-759-143-110 (1-3410) x Q9SP63 (1-501)

Qy        191 AGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAGTGCCTGA 250
              |||:::||||||||||
Db          8 ArgGlnArgGlyArg--------------------------------------------- 12

Qy        251 ACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGTGAGCCGC 310
                    |||  ||||||      |||                |||:::    :::   |||
Db         13 ------ProArgAlaLeuIleGlyGluProValArg---ProArgValProLeuArgArg 29

Qy        311 CTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAG 370
              |||||||||                          |||      :::     |||::::::
Db         30 LeuLeuArg--------------------------ValAlaSerValAlaCysGlyIleGln 41

Qy        371 GTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAG 430
                       |||      ::::::   :::   |||   :::   |||:::|||:::
Db         42 PheGlyTrpAlaLeuGlnLeuSerLeuLeuThrProTyrValGlnGluLeuGlyIlePro 61

Qy        431 GAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCG 490
                 :::   :::::::::            |||||||:::  |||||||:::       |||
Db         62 HisAlaTrpSerSerIleIleTrpLeuCysGlyProLeuSerGlyLeuLeuValGlnPro 81

Qy        491 CTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATC 550
              |||:::|||       ||||||           |||:::|||||||||||||||||||||
Db         82 LeuValGlyHisLeuSerAspArgCysAsnSerArgPheGlyArgArgArgProPheIle 101

Qy        551 TGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTA 610
              |||   :::   |||::::::::::::   :::|||          :::         :::
Db        102 ValAlaGlyAlaThrSerIleValValAlaValLeuIleIleGlyPheSerThrAspIle 121

Qy        611 GCAGGGCTGCTG------TGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTG 664
              ||||||||||||            |||   ||||||         :::|||   ::::::
Db        122 GlyGlyLeuLeuGlyAspGlyAlaAspArgArgProArgAlaValAlaThrPheValVal 141

Qy        665 GGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTC 724
              |||    |||||||||               |||         |||      |||||||||
Db        142 GlyPheTrpLeuLeuAspValAlaAsnAsnValThrGlnGlyProCysArgAlaLeuLeu 161

Qy        725 TCTGACCTCTTCCGGGACCCGGACCAC------TGTCGCCAGGCCTACTCTGTCTATGCC 778
              :::||||||         :::   ||||||         |||   |||   :::   ::::::
Db        162 AlaAspLeu---ThrGluLysAspHisArgArgThrArgValAlaAsnAlaTyrPheSer 180

Qy        779 TTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC--------------- 823
                    |||::::::|||      ||||||:::      :::
Db        181 LeuPheIleAlaValGlyAsnValLeuGlyPheAlaThrGlySerTyrSerGlyTrpPhe 200

Qy        824 ---ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTC 880
                 |||   |||   ||||||         :::                  :::|||
Db        201 ArgIlePheTrpPheThrSerThrSerSer------------CysAsnAlaAspCysAla 216

Qy        881 ------------TTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTG 928
                          ||||||   :::||||||         :::|||   |||
Db        217 AsnLeuLysSerAlaPheLeuLeuAspIleIlePhe---------IleAlaIleThrThr 233

Qy        929 CTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCC 988
                 ::::::   ||||||                         ||||||:::
Db        234 TyrIleSerIleThrAlaAla---------------GlnGluLeuProLeuSerSerSer 248

Qy        989 TCCTTGTCGCCCCACTGCTGT---------------CCATGCCGGGCCCGCTTGGCTTTC 1033
              |||   |||   |||                          |||:::   ||||||
Db        249 SerArgSerThrHisIleSerGluGluMetAlaGluSerThrHisAlaGlnGluAlaPhe 268

Qy       1034 CGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACC 1093
                                        |||   :::|||                  |||
Db        269 --------------------------LeuTrpGluLeuPhe-----------GlyThr 275

Qy       1094 CTGCGCCGG-----------------------CTCTTCGTGGCTGAGCTGTGCAGCTGG 1129
              ||||||                          |||||||||   |||   :::|||
Db        276 LeuArgTyrLeuSerGlySerIleTrpIleIleLeuPheValThrAlaLeu---ThrTrp 294

Qy       1130 ATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAG 1189
              :::                 |||   ||||||   |||||||:::::::|||    :::|||
Db        295 IleGlyTrpPheProPheLeuLeuPheAspThrAspTrpMetGlyArgGluIleTyrGly 314

Qy       1190 GGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGG 1249
              |||   |||                   :::::::|||:::   |||||||||
Db        315 GlyLysPro------------------AsnGluGlyGlnAsnTyrAsnThrGlyValArg 328

Qy       1250 ATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATG 1309
              ||||||:::|||||||||   |||         :::      :::   |||:::::::|||
Db        329 MetGlyAlaLeuGlyLeuMetLeuAsnSerValValLeuGlyIleThrSerValLeuMet 348

Qy       1310 GACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTC 1369
              ::::::|||   ::::::::::|||      |||:::
Db        349 GluLysLeuCysArgLysTrpGlyAlaGlyPheValTrp--------------------- 361

Qy       1370 CCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCC 1429
                                                                  |||
Db        362 ---------------------------------------------GlyLeuSerAsn 365

Qy       1430 GCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTC 1489
                 |||            |||      :::   ||||||            |||   :::
Db        366 IleLeuMetSerLeuCysPheLeuLeuMetLeuIleLeuSer---------AlaValVal 382

Qy       1490 TACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGC 1549
                 |||   :::
Db        383 LysHisMetAsp------------------------------------------------ 386
```

```
Db        257 y---------------LeuProGlyMetLysGlyHisArgGlyPheSerGlyLeuAspGl 272

Qy       1892 AGTAAATGGCGACCAGACCCAGGCCTGCGGCAGACACCATATAGGCAGTGACAGACTGGC 1833
                        ||||||           |||       :::         ||||||
Db        272 yAlaLysGlyAspAlaGlyProAlaGlyProLysGlyGluProGlySerProGly----- 290

Qy       1832 TGAGCTGGACAATGGAGCCCATAAACAGGGATGGGGCCA---------CCTGGGACAGCA 1782
                       :::|||||||||        |||   |||||||||          ||||||
Db        291 -------GluAsnGlyAlaProGlyGln--MetGlyProArgGlyLeuProGlyGluArg 307

Qy       1781 GGAAGGCACTATCCAGGATGGCGAGGTCCAGGCAGATGCCCCGGCCCGGAA--------- 1731
              ||||||      ||||||       ||||||              |||   :::
Db        308 GlyArg------ProGlyAlaProGlyProAlaGlyAlaArgGlyAsnAspGlyAlaThr 325

Qy       1730 -----------CCACCCTGGCCTCGGTGGGCTCACCCACCACCACA------------- 1696
                         ||||||   |||              ||||||
Db        326 GlyAlaAlaGlyProProGlyProThrGlyProAlaGlyProProGlyPheProGlyAla 345

Qy       1695 ---------CGTACGGAGACATCACAG---------------GCAGAGGCCCCGCAGAGC 1660
                         :::   |||                        :::|||    ||||||
Db        346 ValGlyAlaLysGlyGluAlaGlyProGlnGlyProArgGlySerGluGlyProGlnGly 365

Qy       1659 GCGGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGGG---AAG 1603
                 |||                 |||   ||||||   |||             |||
Db        366 ValArgGly---------GluProGlyProProGlyProAlaGlyAlaAlaGlyProAla 382

Qy       1602 GGAGCTCCAGGCTTAGGG------CCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTG 1549
              |||   ||||||              ||||||   |||      ::::::|||   |||
Db        383 GlyAsnProGlyAlaAspGlyGlnProGlyAlaLys-GlyAlaAsnGlyAlaPro----- 400

Qy       1548 CTAGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGTAG 1489
                                    ||||||   ||||||   |||      ||||||   ||
Db        401 ---------------------GlyIle---AlaGlyAlaProGlyPheProGlyAlaAr 412

Qy       1488 AGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCG 1429
               ||||             :::|||    ||||||     |||    ||||||      |
Db        412 gGly-----------ProSerGlyPro-GlnGlyProGlyGlyProProGlyProLysG 428

Qy       1428 GCTGAAGCTGTCACCACGGCCACACTCTGGGACAGGCATGTGGCACCGGCAGCCACAGGG 1369
              ||   |||           |||         |||            |||||| |||   |||
Db        428 lyAsnSerGlyGluProGly---AlaProGlySerLysGlyAspThrGly-AlaLysGly 446

Qy       1368 AAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCTGCACCAGCCGGTCC 1309
              :::                                |||   :::   |||   |||||
Db        447 GluProGlyPro--------------------Val-GlyValGlnGlyProProGlyPr 459

Qy       1308 ATGACCAGAG------AGAAGACCAGGGAGATGGCGCACTGCAGGAACAGCC---CCAGG 1258
              |      |||      :::|||       |||               |||      |||||
Db        459 oAlaGlyGluGluGlyLysArgGlyAlaArgGlyGluProGlyProThrGlyLeuProGl 479

Qy       1257 CTGCCCATCCGAACGC-------------------------CTTCATCATAGTGTCTC 1225
              |    |||    ||||||                                  |||::
Db        479 yProProGlyGluArgGlyGlyProGlySerArgGlyPheProGlyAlaAspGlyValAl 499

Qy       1224 CGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTACAGCCCCTCGCCCACGAAA 1165
              :||||||:::     ||||||                  ||||||    |||           ::
Db        499 aGlyProLysGlyProAlaGlyGluArgGlySerProGlyProAlaGlyProLysGlySe 519

Qy       1164 TCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGCTGCACAGCTCAGCCACGAAG 1105
              :|||
Db        519 rPro--------------------------------------------------------- 520

Qy       1104 AGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACAGCTGGTGCAGCCGGGGAAGCAGGGCG 1045
                                                   ||||||     ||||||||||||
Db        521 ----------------------------GlyGluAlaGlyArgProGlyGluAlaGlyLe 531

Qy       1044 CCCAGGT------------TCCGGAAAGCCAAGCGGGCCCGGCATGGACAGCAGTGGGGC 997
                ||||||             :::|||   |||       ||||||    |||
Db        531 uProGlyAlaLysGlyLeuThrGlySerProGlySerProGlyProAspGly-------- 548

Qy        996 GACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCG----------- 948
                  :::   ||||||                         |||||||||
Db        549 ----LysThrGlyProPro---------------------GlyProAlaGlyGlnAspGl 560

Qy        947 -------CTGCCTCCTCAGCCACCA------------------------------------ 930
                            ||||||   ||||||
Db        560 yArgProGlyProProGlyProProGlyAlaArgGlyGlnAlaGlyValMetGlyPhePr 580

Qy        929 -------------GCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCA 883
                           |||                       |||     |||   ||||||::
Db        580 oGlyProLysGlyAlaAla------------------GlyGluProGlyLysAlaGlyGl 594

Qy        882 AAGAGGCACT------CCTCCTGGGTGCCCAGGT-------------------------- 855
              :||||||           |||||||||      |||
Db        594 uArgGlyValProGlyProProGlyAlaValGlyProAlaGlyLysAspGlyGluAlaGl 614

Qy        854 ----------------------AGGGGGCCAGGGCACTGGTGTCCCAGTCAATGGCAGGC 817
                                    ||||||                  |||       ||
Db        614 yAlaGlnGlyProProGlyProAlaGlyProAlaGlyGluArgGlyGluGlnGlyProAl 634

Qy        816 AGGA-----------------------GGTAGCCCAGGCAGCCCCCAAGACTGATCATG 781
              ||||                        |||       |||    |||   :::
Db        634 aGlySerProGlyPheGlnGlyLeuProGlyProAlaGlyProProGlyGluAlaGlyLy 654

Qy        780 AAGGCATAGACAGAGTAGGCCTGGCGACAGTGGT---CCGGGTCCC------GGAAGAGG 730
                     |||      |||||||||   |||    |||||||||           |||||
Db        654 sProGlyGluGlnGlyValProGlyAspLeuGlyAlaProGlyProSerGlyAlaArgGl 674

Qy        729 TCAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCC 670
              |:::|||   |||   |||   |||                |||        ||||||   ||
Db        674 yGluArg---GlyPheProGlyGluArgGlyValGlnGlyProProGlyProAlaGlyPr 693

Qy        669 ACGCCCAGGATGAGCAGTGCCA-----------------------------GCTCCAGG 640
              ||||         |||                                    ||||||||
Db        693 oArgGlyAlaAsnGlyAlaProGlyAsnAspGlyAlaLysGlyAspAlaGlyAlaProGl 713

Qy        639 GGCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGG 580
              ||||        ||||||:::                :::   :::     ||||||||
Db        713 yAla------ProGlySerGlnGlyAlaProGlyLeuGlnGlyMetProGlyGluArgGl 731

Qy        579 CTCAGCAGGATGCCCAAGGACAGTGCCCAGATGAAGGGCCGGCGGCCATAGCGTCCA 520
              |:::||||||   |||           |||:::          |||
Db        731 yAlaAlaGlyLeuProGly------ProLysGlyAspArgGlyAspAlaGlyProLysGl 749

Qy        519 CGCCAGTGGTCACTGGCTGAGCCTAGGAGCG---GGACACAGA----------------- 480
                |||   |||          |||   |||          |||   |||
Db        749 yAlaAspGly---------SerProGlyLysAspGlyValArgGlyLeuThrGlyProIl 766

Qy        479 -------CCAGGCCCAGCACTGGACCAATGCCCA-------------------------- 453
                       ||||||||||||       |||
Db        766 eGlyProProGlyProAla------GlyAlaProGlyAspLysGlyGluSerGlyProSe 784

Qy        452 ------------------------------GCACCATGGTCATGAACTTCTCCTCTACC 424
                                            ||||||                 |||   ||
Db        784 rGlyProAlaGlyProThrGlyAlaArgGlyAlaProGlyAspArgGlyGluProGlyPr 804

Qy        423 CCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCCAAACACACCTCCAGG 364
              ||||   ||||||       |||                              ||||||||
Db        804 oProGlyProAlaGlyPheAlaGly--------------------------ProProGl 815

Qy        363 C------------CAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGTGCCGC 316
              |                             |||   |||      ||||||    |||   ||
Db        815 yAlaAspGlyGlnProGlyAlaLysGlyGluProGlyAspAlaGlyAlaLysGlyAspAl 835

Qy        315 AGCAGGCGGCTCACCCACAGCCTCTGGACCATAGTGGGCCA---------GGCGGGT 268
              |      |||   |||       :::||||||  :::||||||                |||
Db        835 a------GlyProProGlyProAlaGlyPro-AlaGlyProProGlyProIleGlyAsnV 853

Qy        267 AGGGCTCAGGGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCTCGGCTCTGCT------C 214
                |||:::   |||      |||                      |||||||||        |
Db        853 alGlyAlaProGlyAlaLysGlyAla---------------ArgGlySerAlaGlyProP 868

Qy        213 CAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAACTGCCTAGGAATCAGCCAGGCGCCC 154
              ||   |||          |||             :::::            |||        :
Db        868 roGlyAlaThrGlyPheProGlyAlaAlaGlyArgValGlyProProGlyProSerGlyA 888
```

```
Db       3571 nProGly 3573

RESULT 22
O76045
ID   O76045      PRELIMINARY;      PRT;  1461 AA.
AC   O76045;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Pro alpha 1(I) collagen.
GN   COL1A1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85130970; PubMed=2857713;
RA   Chu M.L., de Wet W., Bernard M., Ramirez F.;
RT   "Fine structural analysis of the human pro-alpha 1 (I) collagen gene.
RT   Promoter structure, AluI repeats, and polymorphic transcripts.";
RL   J. Biol. Chem. 260:2315-2320(1985).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88329734; PubMed=2843432;
RA   D'Alessio M., Bernard M., Pretorius P.J., de Wet W., Ramirez F.;
RT   "Complete nucleotide sequence of the region encompassing the first
RT   twenty-five exons of the human pro alpha 1(I) collagen gene.";
RL   Gene 67:105-115(1988).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89025644; PubMed=3178743;
RA   Tromp G., Kuivaniemi H., Stacey A., Shikata H., Baldwin C.T.,
RA   Jaenisch R., Prockop D.J.;
RT   "Structure of a full-length cDNA clone for the prepro alpha 1(I) chain
RT   of human type I procollagen.";
RL   Biochem. J. 253:919-922(1988).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91138770; PubMed=1995349;
RA   Maatta A., Bornstein P., Penttinen R.P.;
RT   "Highly conserved sequences in the 3'-untranslated region of the
RT   COL1A1 gene bind cell-specific nuclear proteins.";
RL   FEBS Lett. 279:9-13(1991).
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92157916; PubMed=1787829;
RA   Westerhausen A., Constantinou C.D., Pack M., Peng M.Z., Hanning C.,
RA   Olsen A.S., Prockop D.J.;
RT   "Completion of the last half of the structure of the human gene for
RT   the Pro alpha 1 (I) chain of type I procollagen (COL1A1).";
RL   Matrix 11:375-379(1991).
RN   [6]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98107942; PubMed=9443882;
RA   Korkko J., Ala-Kokko L., De Paepe A., Nuytinck L., Earley J.,
RA   Prockop D.J.;
RT   "Analysis of the COL1A1 and COL1A2 genes by PCR amplification and
RT   scanning by conformation-sensitive gel electrophoresis identifies only
RT   COL1A1 mutations in 15 patients with osteogenesis imperfecta type I:
RT   identification of common sequences of null-allele mutations.";
RL   Am. J. Hum. Genet. 62:98-110(1998).
RN   [7]
RP   SEQUENCE FROM N.A.
RA   Korkko J.M., Earley J.J., Nuytinck L., DePaepe A., Prockop D.J.,
RA   Ala-Kokko L.;
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF017178; AAB94054.2; -.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR000885; Fib_collagen_C.
DR   InterPro; IPR001007; VWF_C.
DR   Pfam; PF01410; COLFI; 1.
DR   Pfam; PF01391; Collagen; 18.
DR   Pfam; PF00093; vwc; 1.
DR   ProDom; PD000007; Collagen; 2.
DR   ProDom; PD002078; Fib_collagen_C; 1.
DR   SMART; SM00038; COLFI; 1.
DR   SMART; SM00214; VWC; 1.
DR   PROSITE; PS01208; VWFC; 1.
KW   Collagen.
SQ   SEQUENCE   1461 AA;  138630 MW;  9ACF6DE30EA78E21 CRC64;

  Alignment Scores:
  Pred. No.:              6.3e-14        Length:        1461
  Score:                  342.50         Matches:       283
  Percent Similarity:     31.26%         Conservative:  49
  Best Local Similarity:  26.65%         Mismatches:    346
  Query Match:            5.50%          Indels:        385
  DB:                     4              Gaps:          62

US-09-759-143-110 (1-3410) x O76045 (1-1461)

Qy       2591 CTGTGTCTCTGTGATGGCAACAGA-------------------------AGGACCAAC 2559
              :::|||:::||||||        :::                         |||
Db         60 IleCysValCysAspAsnGlyLysValLeuCysAspAspValIleCysAspGluThr--- 78

Qy       2558 AGGCCACATCCTGATAAAAGGTAAGAGGGGGGTGGATCAGCAAAAAGACAGTGCTGTGGG 2499
                                                             :::  |||  |||
Db         79 ------------------------------------------LysAsnCysProGly 83

Qy       2498 CTGAGGGGACCT---GGTTCTTGTGTGT------TGCCCCTCAGGACTCTTCCCCTACAA 2448
                        |||   |||  ||| |||       ||||||  |||
Db         84 AlaGluValProGluGlyGluCys-CysProValCysProAspGly-------------- 98

Qy       2447 ATAAGTCATATGTTCAAATCCCATGGAGGAGTGTTTCATCCTAGAAACTCCCATGCAAGA 2388
                                                     |||:::|||
Db         99 ------------------------------------SerGluSerPro-------- 102

Qy       2387 GCTACATTAAACGAAGCTGCAGGTTAAGGGGCTT------------------------- 2354
                        |||     :::      |||:::
Db        103 ---------ThrAspGlnGluThrThrGlyValGluGlyAspThrGlyProArgGlyPr 119

Qy       2353 ------------------------AGAGATGGGAAACCAGGTGACTGAGTTTATTCAGC 2319
                                      |||||||||  ||||||
Db        119 oArgGlyProAlaGlyProProGlyArgAspGlyIleProGlyGlnProGlyLeuProGl 139

Qy       2318 TCCCAAAAACCCT------TCTCTAGGTGTGTCTCAACTAGGAGGCTAGCTGTTA---AC 2268
               |||     |||            |||        |||||||||
Db        139 yProProGlyProProGlyProProGlyProProGlyLeuGlyGlyAsnPheAlaProGl 159

Qy       2267 CCTGAGCCTGGGTAATCCACCTGCAGAGTCCCCGCATTCCAGTGCATGGAGCCCTTCTGG 2208
               ||||||  |||                 :::        |||         |||  ||
Db        159 nLeuSerTyrGlyTyrAspGluLysSerThrGlyGlyIleSerValProGlyProMetGl 179

Qy       2207 CCTCCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAG 2148
              |                                             ||||||
Db        179 y-----------------------------------------ProSerGlyPro----- 183

Qy       2147 AGACTGGGGAGAGAGGAGAGGGACGCCCCAGCCCCCAGCTGTGCAGCTACGCACCTCAGC 2088
                                ||||||  |||   ||||||
Db        184 ----------------ArgGlyLeuProGlyProProGlyAlaPro---GlyProGlnGl 197

Qy       2087 AGCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGC 2028
                  ||||||         ||||||       |||::::::          |||
Db        197 yPheGlnGlyProProGlyGluProGlyGluProGlyAlaSerGlyProMetGlyProAr 217

Qy       2027 CCGGCAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGGA------------------- 1988
                    |||         ::::::|||:::   |||
Db        217 gGlyProProGlyProProGlyLysAsnGlyAspAspGlyGluAlaGlyLysProGlyAr 237

Qy       1987 -CCCAGTGAGGCAGGCCCTCCA-----------------------CCCCAATGTGCTGG 1953
               |||   |||   |||||||||                         |||    |||||
Db        237 gProGlyGluArgGlyProProGlyProGlnGlyAlaArgGlyLeuProGlyThrAlaGl 257

Qy       1952 AAGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAA 1893
              |                 :::  |||                         :::
```

```
Qy      1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
                                |||
Db      3024 ------------GluTrp--------------------------------------- 3025

Qy      1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
                 ||||||            |||           |||||||||||     :::|||    :
Db      3026 ----GlyProTrpSerPro-CysSerValProCysGlyGlyGlyTyrArgAsnArgThrA 3044

Qy      1724 AGGGTGGTTCCGGGCCGGGGCATCTG----------------------------CCTG 1753
             ::|||   |||||||||          :::                          |
Db      3044 rgGlySerSerGlyProSerProValAspPheSerThrCysGlyLeuGlnProCysAlaG 3064

Qy      1754 GACCTCGCCATCCTGGATAGTGCCTTCCT------------------------------ 1782
             |||||        ||||||   |||    |||
Db      3064 lyProAla---ProGlyValCysProProGlyLysArgTrpLeuAspCysAlaGlnGlyP 3083

Qy      1783 ----------------------GCTGTCCCAGGTGGCCCCATCCCTGTTTAT---GGGC 1816
                                   |||        |||        |||    :::    |||
Db      3083 roAlaSerCysAlaGluLeuSerAlaProArgGlyAlaAspGlnProCysHisProGlyC 3103

Qy      1817 TCCATTGTCCAGCTCAGCCAGTC-------------------------------TGTC 1843
               :::|||||||:::      :::                                |||
Db      3103 ysTyrCysProSerGlyMetLeuLeuLeuAsnAsnAlaCysValProThrGlnAspCysP 3123

Qy      1844 ACTGCCTATAT---GGTGTCTGCCGCAGGCCTGGGTCTGGT--------------CGCC 1885
               |||     :::    |||          ||||||||                   :
Db      3123 roCysThrHisGlyGlyArgLeuHisProProGlySerAlaValLeuArgProCysGluA 3143

Qy      1886 ATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACT---TGGCCAAATACTCAGCG 1942
             ::          |||    :::|||    |||   |||          |||    ||||||     :::
Db      3143 snCysSerCysValSerGlyLeuIle--ThrAsnCysThrSerTrpProCysLysGluGl 3162

Qy      1943 TAGAAAACTTCCAGCACATTGG------------------------------------- 1964
               :::          :::     |||
Db      3162 yGlnProThrTrpSerProTrpThrProTrpSerGluCysSerAlaSerCysGlyProAl 3182

Qy      1965 ---------------------------------------GGTGGAGGGCCTGCCTCACTGGG 1987
                                                   ||||||      |||:::|||:::
Db      3182 aArgArgHisLysHisArgPheCysThrArgProProGlyGlyAlaProSerSerMetAl 3202

Qy      1988 TCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTT 2047
               |||          ||||||||||||:::
Db      3202 aPro------ProLeuLeuLeuSerSer----------------------------- 3209

Qy      2048 GCTGCCAAAGTAATGTGGCTCT-----------------------CTGCTGCCACCCTG 2083
               :::|||        |||                          |||||||||    ||
Db      3210 -ValProProLeuCysProGlyProGluAlaGluGluGluProCysLeuLeuProGluCy 3229

Qy      2084 TGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCA 2143
             |                                  ||||||||||||||
Db      3229 sAsp-----------------ArgAlaGlyGlyTrpGly------------------- 3236

Qy      2144 GTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGG 2203
                                     |||     |||                   |||:::
Db      3237 ---------------ProTrpGlyPro-----------------TrpSerSerCysSe 3245

Qy      2204 GAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTA---CCCAG 2260
               |||                             ||||||          |||||||||    ||
Db      3245 rArg--------------------SerCysGly---------GlyGlyLeuArgSerAr 3255

Qy      2261 GCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGA---------------- 2303
             |   |||                ||||||               |||:::
Db      3255 gSerArgAlaCysAspGlnProPro-----------ProGlnGlyLeuGlyAspTyrCy 3271

Qy      2304 -GAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTT 2362
              ||||||                                    |||
Db      3271 sGluGly-----------------------------------ProArgAlaGlnGlyAl 3279

Qy      2363 AACCTGCAGCTTCGTTTAATGT------AGCTCTTGCATGGGAGTTTCTAGGATGAAACA 2416
               |||        :::    |||          ::::::|||         :::          :::::
Db      3279 aAlaCysGlnAlaLeuProCysProValThrAsnCysThrAlaIleGluGlyAlaGluTy 3299

Qy      2417 CTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACA 2476
               :|||:::    |||
Db      3299 rSerAlaCysGly------------------------------------------- 3303

Qy      2477 CACAAGAACCAGGTCCCCTCAGCCCACAG-----------------CACTGTCTTTTT- 2517
                                  |||    |||:::                  ||||||:::::
Db      3304 ------------ProPro-CysProArgSerCysAspAspLeuValHisCysValTrpH 3319

Qy      2518 --GCTGATCCA-----------CCCCCTCTTACCTTTTATCAGGATGTGGCTGTTGG 2563
                    |||               ||||||       :::||||||       |||    :::
Db      3319 isCysGlnProGlyCysTyrCysProProGlyGlnValLeuSerAlaAspGlyThrValH 3339

Qy      2564 TCCTT-------------------------CTGTTGCCATCACAGAGACACAGG---- 2592
               :::                          ||||||       :::|||||||||
Db      3339 isValGlnProGlyHisCysSerCysLeuAspLeuLeuThrGlyGluArgHisArgProG 3359

Qy      2593 ----------------------CATTTAAATATTTAACTTATTTATTTAACAAAGTAGA 2629
                                    |||          |||   |||       :
Db      3359 lyAlaGlnLeuAlaLysProAspGlyCysAsn-----------TyrCysThrCysSerG 3375

Qy      2630 AGGGA---ATCCATTGC---TAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGG 2683
             ::|||    :::    |||        |||   |||    |||        |||     |||
Db      3375 luGlyGlnLeuThrCysThrAspLeuProCysPro-ValProGlyAlaTrpCysProTrp 3394

Qy      2683 ------------------------------------------------------------ 2683

Db      3395 SerGluTrpThrAlaCysSerGlnProCysGlnGlyGlnThrArgThrArgSerArgAla 3414

Qy      2684 --------------------------GGGATCCCCAACAATCAGGTCCCCTGAGATAGC 2716
                                       |||   |||            |||  ::::::
Db      3415 CysSerCysProAlaProGlnHisGlyGlyAlaPro---------CysProGlyGluAla 3431

Qy      2717 TGGTCATTGGGC----------------------TGATCATTGCCAGAATCTTCTTCT 2752
                      |||                         |||    ||||||
Db      3432 GlyGluAlaGlyAlaGlnHisGlnArgGluThrCysAlaSerThrProGluCysProVal 3451

Qy      2753 CCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAA 2812
               |||    |||    |||                 |||||||||
Db      3452 AspGlyAlaTrpSerProTrp-----------GlyProTrpSer-------------- 3462

Qy      2813 ATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGG 2872
                      |||       |||       |||  |||
Db      3463 ---------ProCysGluValCysLeuGly-Arg----------------------Se 3471

Qy      2873 CTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCC 2932
                  ||||||:::                     ||||||              |||||
Db      3471 rHisArgSerArgGluCys-----------------SerTrp-----------ProPr 3481

Qy      2933 CACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTT 2992
              |||||||                                  ||||||
Db      3481 oThrSer---------------------------------GluGlyGly---------- 3486

Qy      2993 CCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTG 3052
                                      |||    |||        |||:::         |||
Db      3487 ---------------ArgProCysProGly---GlyHisArgGlnSerArgPro----- 3498

Qy      3053 GAGCTACTGCAGGACCAGAAGCACAAAGTGC---------------------------- 3083
                     |||:::        ||||||:::|||
Db      3499 -------CysGlnGlyAsnSerThrGlnCysThrAspCysAlaGlyGlyGlnAspLeuLe 3516

Qy      3084 ------------GGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGT-- 3128
                                |||       |||                |||           |||
Db      3516 uProCysGlyGlnProCysProArgSerCysGluAspLeuSerProGlyValGluCysGl 3536

Qy      3129 -------GCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGG 3181
                       :::    |||                   |||    ||||||
Db      3536 nProAspSerMetGlyCysGlnGlnProArgCysGlyCysProPro------------- 3551

Qy      3182 AGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTT 3241
               |||    :::||||||    |||                   ||||||       |||
Db      3552 -GlyGlnLeuSerGlnAspGlyLeuCysValThrProSerGlnCysArgCysGlnTyrGl 3571

Qy      3242 AGCGGGG 3248
                |||
```

```
RP   SEQUENCE FROM N.A.
RC   TISSUE=SUBCOMMISSURAL ORGAN;
RA   Meiniel A.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SUBCOMMISSURAL ORGAN;
RX   MEDLINE=20465125; PubMed=11008217;
RA   Gobron S.;
RT   "Subcommissural organ/Reissner's fiber complex: characterization of
RT   SCO-spondin, a glycoprotein with potent activity on neurite
RT   outgrowth.";
RL   Glia 32:177-191(2000).
DR   EMBL; AJ416457; CAC94914.1; -.
SQ   SEQUENCE   5146 AA;  543576 MW;  724C5FB8727E13DA CRC64;

Alignment Scores:
Pred. No.:              5.76e-14      Length:         5146
Score:                  344.50        Matches:        308
Percent Similarity:     29.12%        Conservative:   79
Best Local Similarity:  23.18%        Mismatches:     379
Query Match:            5.37%         Indels:         564
DB:                     6             Gaps:           87

US-09-759-143-110 (1-3410) x Q8SPM4 (1-5146)

Qy    120 GAAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGT 179
Db   2529 GlnAlaCysProValAlaGlyAlaTrpAlaGluTrpGlu------------------Ala 2542

Qy    180 TGGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCT---------GGAGCAGAGCCGAGACGA 230
Db   2543 Trp-----------GlyProCysSerValSerCysGlyGlyGlyHisArgSerArgArg 2558

Qy    231 AGCAGTTCTGGAGTGCCTGAACGGCCCCCTGAG------------CCCTACCCGCCTGGC 278
Db   2559 ArgSerCysMetAspPro------ProProLysAsnGlyGlyAlaProCysProGlyPro 2576

Qy    279 CCACTATGGTCCAGAGGCTGTGGGTGAGCCGCCTGC---------------TGCGGCACC 323
Db   2577 ProGlnGluArgAlaProCysGlyLeuGlnProCysAlaGlyGlyThrAspCysGlyGln 2596

Qy    324 GGAAAGCCCAGCTCTTGCTGGTCAACCTGC---------------TAACCTTTGGCCTGG 368
Db   2597 GlyArgValHisValSerAlaGluLeuCysArgLysGlyLeuValProProCysProPro 2616

Qy    369 AGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAG 428
Db   2617 SerCysLeuAspProGluAlaAsnArgSerCysSerGlyLeuCysLeuGlu-GlyCysAr 2636

Qy    429 AGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCC 488
Db   2636 g-------------------------------------------------------CysPr 2638

Qy    489 CGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCA 548
Db   2638 oProGlyLeuLeuLeuGlnAspAlaGlyCysLeuProLeuSerGluCysProCysLeu-- 2657

Qy    549 TCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGC 608
Db   2658 -ValGlyGluGluLeuGlnGlnProGlyValPro-PheLeuLeuAspAsnCysSerArgC 2677

Qy    609 TA------------GCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGC 656
Db   2677 ysValCysGluLysGlyAlaLeuLeuCys---Glu-ProGlyGlyCysProValProCys 2695

Qy    657 TCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGG 716
Db   2696 ---------GlyTrpSerAlaTrpSerSer------------------------TrpGly 2704

Qy    717 CCCTGCTCTCTGACC---TCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCT 773
Db   2705 ProCysAspArgSerCysGlySerGlyLeuArgAlaArgPheArgSerPro------Ser 2722

Qy    774 ATGCCTTCATGATCAGTCTTGGGGGCTGCCTGG------------------GCTACCTCC 815
Db   2723 AsnProProAlaAlaSerGlyGlyAlaProCysGluGlyGlnArgGlnGluLeuGlnAla 2742

Qy    816 TGCCTGCCATTGACTGGGACACCAGTGCCC---TGGCCCCCTACCTGGGCACCCAGGAGG 872
Db   2743 CysTyrSerAlaCysGlyAlaGluValProGlyTrpThrPro---TrpAlaProTrpSer 2761

Qy    873 AGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGG 932
Db   2762 AlaCysSerGlnSerCysLeuValProGlyGlyGlyProAla------------------ 2775

Qy    933 TGGCTGAGG---AGGCAGCGCTGGGCCCCA----------------------------- 959
Db   2776 ---LeuArgSerArgSerArgLeuCysProGlyProGlyAspThrSerCysIleGlyGlu 2794

Qy    960 CCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGG 1019
Db   2795 AlaThrGluGluGluProCysSerProProValCysLeuGlyLeuGlyVal--------- 2811

Qy   1020 CCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCC 1079
Db   2812 ------Trp------GlyGlnTrpAla-----------------------AlaTrpSer 2819

Qy   1080 GCATGCCCCGCACCCTGCGCCGGC---------TCTTCGTGGCTGAGCTGTGCAGCT--- 1127
Db   2820 AlaCysSerAlaProCysAsnGlyGlyValGlnThrArgGlyArgArgCysSerAlaSer 2839

Qy   1128 ----------------------GGATGGCACTCATGACCTTCACGGTGTTTACACGG 1163
Db   2840 AlaProGlyAspProGlyCysGlnGlyProHisSerGlnThrArgAspCysAsnThr--- 2858

Qy   1164 ATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCC 1223
Db   2859 ------------GlnProCysThrAlaGlnCys-----------------------Pro 2866

Qy   1224 GGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCA 1283
Db   2867 GlyAspMetValPheArgSerAlaGluGlnCysArgTrpGlu---------GlyGlyPro 2883

Qy   1284 TCTCCCTGGTCTTCTCTCTGGTCATGGACCG--------------------------GC 1316
Db   2884 -CysProGlyLeuCysLeuAlaArgGlyProGlyValGluCysThrGlyValCysThrAl 2903

Qy   1317 TGGTGCAGCG---------------ATTCGGCACTCGAGCAGTCTATTT------GGCCA 1355
Db   2903 aGlyCysAlaCysProThrGlyLeuPheLeuHisAsnSerSerCysLeuProProSerGl 2923

Qy   1356 GTGTGGCAGCTTTCCCTGT--------------------------------------- 1374
Db   2923 nCys---------ProCysGlnLeuArgGlyGlnLeuTyrAlaProGlyAlaValAlaAr 2940

Qy   1375 ----------------GGCTGCCGGTGCCA------------------------CATGC 1393
Db   2940 gLeuAspSerCysSerAsnCysThrCysIleSerGlyGluMetValCysAlaSerGluPr 2960

Qy   1394 CTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCA 1453
Db   2960 oCysProValAlaCysGlyTrp------------------SerProTrpThrProTrpSe 2974

Qy   1454 GCCCTGC---AGATCCTGCCCTACA--------------------------CACTGGCC 1483
Db   2974 rLeuCysSerArgSerCysAsnValGlyValArgArgArgPheArgAlaGlyThrAlaPr 2994

Qy   1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
Db   2994 oProAlaAlaPheGlyGlyAlaAlaCysGlnGlyProAsnMetGluAla------GluPh 3012

Qy   1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
Db   3012 eCysSerLeuArgProCysGlyGlyProAlaGly------------------------- 3023
```

```
                                       ·:::|||  |||      :::   |||:::||
Db        933 -------------------------LeuProGlyProAlaGlyGluLysGlyGlyLysGl 944

Qy       1416 TGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACA 1475
              |:::
Db        944 yGlu-------------------------------------------------------- 945

Qy       1476 CACTGGCCTCCCTCTACCACCG-------------------------------------- 1497
                  |||||||||  ||||||
Db        946 -ProGlyLeuProGlyProProGlyProMetHisProAspPheLeuGlySerLysGlyGl 965

Qy       1498 ----GGAGAAGCAGGTGTTCCTGCC---------------------------------CA 1520
                  ||||||  |||:::|||
Db        965 uLysGlyGluProGlyLeuProGlyIleProGlyLeuSerGlyProLysGlyTyrGlnGl 985

Qy       1521 AATACCGAGGGGACACTGGA------GGTGCTAGCAGTGAGGACAGCCTGATGACCAGCT 1574
                :::||| ||||||  |||      |||  |||  :::
Db        985 yLeuPro-GlyAspProGlyGlnProGlyLeuSerGlyGln------------------- 998

Qy       1575 TCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCA 1634
                   ||||||     ||||||  |||  |||  |||:::  |||
Db        999 -----ProGly---SerProGlyLeuProGlyProLysGlyAsnProGly---------- 1012

Qy       1635 GTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTAC 1694
                        ||||||  ||||||  |||  |||              ::::::
Db       1013 --------LeuProGlyProProGlyLeuThrGlyProProGlyProLysGlyAsnIleG 1030

Qy       1695 GTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGG 1754
                   :::|||  |||        ||||||   ||||||         |||||
Db       1030 ly---AspMetGlyPheProGlyPro--GlnGlyValAspGlyProProGlyProProGl 1048

Qy       1755 A---CCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGT---------------GG 1796
              |   |||   ||||||         |||  :::||||||                  ||
Db       1048 yPheProGlyGlnProGly------SerProGlyLeuProGlyGlnLysGlySerLysGl 1066

Qy       1797 CCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGG 1856
              |         |||  |||:::    |||     |||                    ||
Db       1066 yGluProGlyValSerGlyIleGlyLeuProGlyLeuPro------------------Gl 1080

Qy       1857 TGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACA 1916
              |          ||||||                |||    |||  |||:::
Db       1080 yProLysGlyGluProGly------------LeuProGlyTyrProGlyAsn-------- 1093

Qy       1917 AGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCT 1976
                                   ::::::          :::|||     |||||
Db       1094 ------------------ProGlyIleLysGly------SerValGlyGluThrGlyLe 1105

Qy       1977 GCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGC 2036
              ||||                |||                ||||||||||||:::|||
Db       1105 uProGlyLeu-ProGlyAsnProGlyAlaLysGlyGlnProGlyLeuProGlyPheProG 1125

Qy       2037 CAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCC-----------TGT 2084
                          ||||||       |||      ||||||
Db       1125 lyThrProGly-LeuProGlyProLysGlyMetAsnGlyProProGlyAsnProGlyPhe 1144

Qy       2085 GCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAG 2144
                    |||                |||||||||        |||        |||
Db       1145 ProGlyGluProGlyProValGlyAlaGlyGlyArgProGlyProProGlyProProGly 1164

Qy       2145 TCTCTAGGGCTGCCT------------------------------------------GAC 2162
                    |||   |||
Db       1165 GluLysGlyLysProGlyGlnAspGlyIleProGlyProAlaGlyGlnLysGlyGluPro 1184

Qy       2163 TGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCAT 2222
                    :::|||     |||:::                         |||  ||||||
Db       1185 GlyGlnProGlyPheGlyIlePro------------------GlyProProGlyLeuPro 1198

Qy       2223 GCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCT 2282
                 |||                ||||||  |||  |||     |||
Db       1199 GlyLeuSerGlyGlnLysGlyAspGlyGlySerProGlyLeuProGlyAsn--------- 1215

Qy       2283 CCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGT 2342
                                |||      |||::::::              ||||||
Db       1216 ------------------ProGlyLeuProGlySerLys------GlyGluProGly 1226

Qy       2343 TTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTT 2402
              |||                 ::: |||
Db       1227 PheGlnGlyPheProGlyMet-Pro----------------------------------- 1234

Qy       2403 TCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTAGGGGAAGAGT 2462
                    |||        ||||||                         |||:::
Db       1235 ----GlyProProGlyProProGlySer---------------------ProGlyGlnAl 1246

Qy       2463 CCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGA 2522
                 |||   |||             |||||||||    |||
Db       1246 aLeuGluGlyProLysGlyAsnProGlyProGlnGlyProPro----------------- 1260

Qy       2523 TCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACA 2582

Db       1260 ------------------------------------------------------------ 1260

Qy       2583 GAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTG 2642

Db       1260 ------------------------------------------------------------ 1260

Qy       2643 CTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGATCCCCAACAATCAGG 2702
                                              ||||||  |||  |||        ||
Db       1261 -------------------------------GlyArgProGlyProProGlyPheGlnGl 1270

Qy       2703 TCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCT 2762
              ||||
Db       1270 yPro-------------------------------------------------------- 1271

Qy       2763 GGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTC 2822
                  |||        ||||||                        |||      |||
Db       1272 ----ProGlyThrGluGlyProArgGly------------------ProProGlyAsn-- 1283

Qy       2823 CAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCT 2882
                                                         ||||||    |||
Db       1284 -------------------------------------------GlyGlyIleLysGlyGl 1289

Qy       2883 CAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCA-- 2940
                 |||  |||                       ||||||    |||  ||||||
Db       1289 uArgGlyAsnPro------------------GlyProProGlyGlnProGlyLeuProGl 1303

Qy       2941 ----------------------------CTCCCCTCTACTCTCTCTAGGACTGGGCTG-- 2970
                                          ||||||     |||   ||||||
Db       1303 yLeuLysGlyAspGlnGlyProProGlyLeuProGlyAsnProGlyArgProGlyLeuAs 1323

Qy       2971 ----ATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCC-------------- 3012
                  ||||||    |||   :::             |||  ||||||
Db       1323 nGlyMetLysGlyAspProGlyLeuProGlyValProGlyPheProGlyMetLysGlyPr 1343

Qy       3013 -------CTACCC---CCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGA-------- 3054
                     :::|||     |||   |||                 |||   |||
Db       1343 oIleGlyValProGlySerThrGlyProAspGlyGluProGlyLeuThrGlyProProGl 1363

Qy       3055 -------GCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 3087
                            ||||||     ||||||   |||
Db       1363 yProProGlyLeuProGlyProSerGlyGlnSerIleVal 1376
```

```
RESULT 21
Q8SPM4
ID   Q8SPM4      PRELIMINARY;      PRT;   5146 AA.
AC   Q8SPM4;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   SCO-spondin.
GN   SCO-SPONDIN.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
```

```
Db        489 SerGlyValLeuAlaIleValLeuLeuProLysProSerLysAspAlaAlaSerLysLeu 508

Qy       1682 GATGTCTCC 1690
                 :::|||
Db        509 SerLeuSer 511

RESULT 20
Q9ESQ2
ID   Q9ESQ2      PRELIMINARY;      PRT;  1691 AA.
AC   Q9ESQ2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Type IV collagen alpha 5 chain.
GN   COL4A5.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20536494; PubMed=10965041;
RA   Saito K., Naito I., Seki T., Oohashi T., Kimura E., Momota R.,
RA   Kishiro Y., Sado Y., Yoshioka H., Ninomiya Y.;
RT   "Differential Expression of Mouse a5(IV) and a6(IV) Collagen Genes in
RT   Epithelial Basement Membranes.";
RL   J. Biochem. 128:427-434(2000).
DR   EMBL; AB041350; BAB13673.1; -.
DR   MGD; MGI:88456; Col4a5.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR001442; ProcollagnC4.
DR   Pfam; PF01413; C4; 2.
DR   Pfam; PF01391; Collagen; 24.
DR   ProDom; PD000007; Collagen; 3.
DR   ProDom; PD003923; ProcollagnC4; 1.
DR   SMART; SM00111; C4; 2.
KW   Collagen.
SQ   SEQUENCE   1691 AA;  161823 MW;  81340DF1792208FA CRC64;

Alignment Scores:
Pred. No.:              4.14e-14       Length:         1691
Score:                  345.50         Matches:        274
Percent Similarity:     28.91%         Conservative:   55
Best Local Similarity:  24.08%         Mismatches:     369
Query Match:            5.38%          Indels:         440
DB:                     11             Gaps:           57

US-09-759-143-110 (1-3410) x Q9ESQ2 (1-1691)

Qy        129 CCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCAGTTGGCGGCAG 188
              |||  |||  |||              |||||||||   |||   ||||||
Db        532 ProGlyProProGly------------AlaProGly---PheProGlySerLysGlyAsp 546

Qy        189 CAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAGTGCCT 248
                 |||                     |||  :::        :::         |||  |||
Db        547 ProGlyAspValLeuThrLeuProGlyMetLysGlyAspLysGlyGluLeuGlyPhePro 566

Qy        249 GAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGT-------------------- 288
                    |||       |||       ||||||      |||
Db        567 GlyAlaProGlyLeuProGlyLeuProGly-SerProGlyLysAspGlyLeuProGlyLe 586

Qy        289 -CCAGAGGCTGTGGGTGAGCCG--------------------CCTGCTGCGGCACCGGA 326
               |||      |||||||||                                   |||
Db        586 uProGlyProLysGlyGluProGlyGlyIleThrPheLysGlyGluArgGlyProProGl 606

Qy        327 AAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAG 386
                ||||||  |||   |||   |||   |||:::   |||   |||
Db        606 ySerProGlyLeuProGlyLeuProGlyAsnMetGlyProThrGly-------------- 621

Qy        387 GCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCA 446
                                            |||   |||      |||   :::    :::
Db        622 ---------------------ProValGlyPheGlyProProGlyProIleGlyGluLy 634

Qy        447 TGGT-----------GCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCC 494
              |||            ||||||       :::        |||||||||   |||   ||
Db        634 sGlyIleGlnGlyValAlaGly------AsnProGlyGlnProGlyLeu---ProGlyPr 651

Qy        495 TAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGG 554
              |:::
Db        651 oLys------------------------------------------------------- 652

Qy        555 CACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAG 614
                        |||   |||   :::   ::::::   |||                   ::
Db        653 ----------GlyAspProGlyGlnThrIleThrGlnPro--------------GlyLy 664

Qy        615 GGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGG-------- 666
              :        :::|||||||||                           ||||||
Db        664 sProGlyLeuProGlySerProGlyArgAspGlyGluValGlyLeuProGlyAspProGl 684

Qy        667 ----CGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGC 722
                    |||      ||||||         ||||||:::|||   |||   |||   |||
Db        684 yLeuProGlyGlnProGlyLeu------ProGlyIleLeuGlySerLysGlyGluProGl 702

Qy        723 TCTCTGACCTCTTCCGGGACCCGGACCACTGTCG------------CCAGGCCTACTCTG 770
                :::        :::      |||||||||               ||||||:::
Db        702 yIleProGlyIleGlyProProGlyProProGlyProLysGlyPheProGlyIlePro-- 721

Qy        771 TCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACT 830
                                         |||   ||||||   |||   |||           ::
Db        722 ------------------------GlyProProGlyAlaProGlyAlaProGlyArgMe 733

Qy        831 GGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGA-------------------- 870
              :|||       |||         ||||||   ||||||
Db        733 tGlyProAspGlyProProGlyProProGlyPheProGlyProLysGlyGluProGlyPh 753

Qy        871 -GGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGC 929
               |||:::|||      |||      |||      |||      :::   :::
Db        753 eGlyLeuProGlyProProGlyProProGlyPheProGlyPheLysGlyThrProGlyPr 773

Qy        930 TGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCT 989
                 |||      |||      ||||||                       ||||||
Db        773 oLysGlyAspArgGlyPheProGlyProSer--------------------GlyProPr 786

Qy        990 CCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCC 1049
                       |||      :::|||||||||   |||         |||
Db        786 oGlyArgThrGlyLeuAspGlyLeuProGlyProLysGlyAspValGlyProAsnGlyGl 806

Qy       1050 TGCTTCCCCGGCTGCACCAGCTGTGCTGCCG------------CATGCCCCGCACCCTGC 1097
                 |||      |||      ||||||                        |||
Db        806 nProGlyProValGlyProProGlyLeuProGlyIleGlyLeuGlnGlyProProGlyPr 826

Qy       1098 GCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTT-- 1155
                |||      |||                           |||      ||||||   :::
Db        826 oProGlyThrProGlyProIle------------GlyGlnProGlyLeuHisGlyIlePr 842

Qy       1156 ----------------------------------------------TTACACGGATTTCG 1169
                                                            |||.  |||   ||
Db        842 oGlyGluLysGlyAspProGlyProProGlyPheAspValProGlyLeuProGlyGluAr 862

Qy       1170 TGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGAC 1229
               ||||          :::||||||                     |||   ||||||
Db        862 gGlySerProGlyLeuProGlyAla------------ProGlyLeuIleGlyProPro-- 877

Qy       1230 ACTATGATGAAGGCGTTCGGATGGGCAGCCTGGG---------------GCTGTTCCTGC 1274
                                  |||   |||   ||||||                   |||
Db        878 -------------GlySerProGlyValProGlyLysAlaGlyValProGlyPheProGl 893

Qy       1275 AGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGT-------------- 1320
               |||:::            |||::::::         ||||||   |||
Db        893 yValLysGlyGluMetGlyMetMetGlyProProGlyProProGlyProLeuGlyIlePr 913

Qy       1321 ----------------GCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAG 1364
                                    ::::::            :::   ||||||   |||
Db        913 oGlyArgSerGlyAlaProGlyLeuLysGlyAspAspGlyMetGlnGlyGlnProGly-- 932

Qy       1365 CTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCA---------CAGTGTGGCCGTGG 1415
```

```
Db        790 ---HisProProProGluArgGlySerGlyProArgGlyThrArgProPro-AlaProSe 808

Qy        100 TGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTCATCACTCAGATCCTGGCCGAGGCGC 41
              :::      |||      :::                 |||      |||   ||||||::
Db        808 rSerArgThrGlyGlnArgThrArgGlnArgProGlyHisProProProGluArgGlySe 828

Qy         40 G 40
              :
Db        828 r 828

RESULT 19
Q9SLN7
ID   Q9SLN7      PRELIMINARY;      PRT;   515 AA.
AC   Q9SLN7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Sucrose transporter protein.
GN   CSUT.
OS   Daucus carota (Carrot).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
OX   NCBI_TaxID=4039;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=ROOT;
RA   Wu N., Diao F., Zhang L., Huang M.;
RT   "Sucrose transporter protein.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR   EMBL; AB036758; BAA89458.1; -.
DR   InterPro; IPR003662; sub_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   TIGRFAMs; TIGR01301; GPH_sucrose; 1.
KW   Transmembrane.
SQ   SEQUENCE   515 AA;  54523 MW;  2B548D1D9DFF51AC CRC64;

Alignment Scores:
Pred. No.:              2.93e-14      Length:         515
Score:                  346.50        Matches:        137
Percent Similarity:     42.35%        Conservative:   76
Best Local Similarity:  27.24%        Mismatches:     215
Query Match:            5.40%         Indels:         75
DB:                     10            Gaps:           19

US-09-759-143-110 (1-3410) x Q9SLN7 (1-515)

Qy        332 CAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATC 391
              :::|||:::||||||      :::      |||:::::::          |||     :::
Db         34 LysLeuValLeuValAlaAlaIleAlaAlaGlyValGlnPheGlyTrpAlaLeuGlnLeu 53

Qy        392 ACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTG 451
              :::   :::   |||   :::      :::|||:::      |||:::            :::
Db         54 SerLeuLeuThrProTyrValGlnLeuLeuGlyIleProHisLysTrpAlaAlaTyrIle 73

Qy        452 CTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGAC 511
                          ||||||:::   |||::::::      |||:::::::|||      ||||||
Db         74 TrpLeuCysGlyProIleSerGlyMetLeuValGlnProIleValGlyTyrTyrSerAsp 93

Qy        512 CACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATC 571
              |||   :::      :::|||||||||||||||||||||   :::   :::      :::
Db         94 HisCysGlnSerSerPheGlyArgArgArgProPheIleAlaSerGlyAlaGlyCysVal 113

Qy        572 CTGCTGAGCCTCTTTCTCATCCCAAGGGCC---------GGCTGGCTAGCAGGG---CTG 619
              :::|||:::   ||||||      |||               |||:::   ||||||
Db        114 AlaIleSerValIleLeuIleGlyPheAlaAlaAspIleGlyTyrLysAlaGlyAspAsp 133

Qy        620 CTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTG 679
              :::         :::|||      :::   :::   :::::::|||         :::|||
Db        134 MetSerLysThrLeuLysProArgAlaValThrValPheValIleGlyPheTrpIleLeu 153

Qy        680 GACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGG 739
              |||            :::         |||      |||||||||:::||||||
Db        154 AspValAlaAsnAsnMetLeuGlnGlyProCysArgAlaLeuLeuAlaAspLeuCysSer 173

Qy        740 ---GACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGG 796
                 |||            |||   |||   :::   |||:::|||   :::::::::|||
Db        174 GlyAspThrArgArgMetArgSerAlaAsnAlaPheTyrSerPhePheMetAlaValGly 193

Qy        797 GGCTGCCTGGGCTAC---------------------------CTCCTGCCTGCCATT 826
                 |||||||||                                    |||   |||
Db        194 AsnIleLeuGlyTyrAlaAlaGlySerTyrAsnAsnLeuTyrLysLeuProPheSer 213

Qy        827 GACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGC---CTCTTT 883
                       |||   |||         |||                :::   |||   :::
Db        214 ------LysThrHisAlaCysAspLeuTyrCysAlaAsnLeuLysSerCysPheIleIle 231

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
                 :::      ||||||   :::|||   ||||||      |||   |||      |||
Db        232 SerIleAlaLeuLeuIleIleIleThrValValAla------LeuSerValValArgGlu 249

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
                       |||   :::   |||:::                  |||
Db        250 LysGlnTrpSerPro---AspAspAlaAspAlaAlaAspGluPro--------------- 263

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
                    |||      :::::               |||   ||||||      |||
Db        264 ------ProSerSerGlyLysIleProVal------PheGlyGluLeuLeuGlyAlaLeu 279

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
                 |||                  ||||||   :::      |||   :::
Db        280 LysAspLeu------------ProArgProMetLeuLeuLeuIleValThrCysLeu 295

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              :::|||:::|||         |||   ||||||   ||||||::::::|||         :::
Db        296 AsnTrpIleAlaTrpPheProPheIleLeuPheAspThrAspTrpMetGlyArgGluIle 315

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              |||   |||         |||      |||            :::   ||||||:::|||
Db        316 TyrGlyGlyThr------AlaGlyLysGly-----------LysLeuTyrAspGlnGly 329

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||   |||||||||||||||   |||         :::      :::   |||:::
Db        330 ValArgAlaGlySerLeuGlyLeuLeuLeuAsnSerValValLeuGlyLeuThrSerIle 349

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
                 ::::::   ||||||:::      |||         :::   |||      |||
Db        350 AlaValGluTyrLeuValArgGlyValGlyGlyValLysIle---LeuTrpGlyValVal 368

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTG-------------------------- 1396
                 |||   :::|||   |||
Db        369 AsnPheIleLeuAlaIleGlyLeuValMetThrValValValSerLysValAlaGlnHis 388

Qy       1397 ------TCCCACAGTGTG----------------------GCCGTGGTGACAGCTTCA 1426
                       ||||||                               |||   |||   |||
Db        389 GlnArgGluHisSerAlaAsnGlyGlnLeuLeuProProSerAlaGlyValLysAlaGly 408

Qy       1427 GCC-----------GCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTAC 1474
              |||            ::::::   |||         |||            :::|||:::
Db        409 AlaLeuSerLeuPheSerIleLeuGlyIleProLeuSerIleThrTyrSerIleProPhe 428

Qy       1475 ACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGAC 1534
                 |||||||||:::|||                                    |||
Db        429 AlaLeuAlaSerIleTyrSerSerGlySerGlyAlaGlyGlnGlyLeuSerLeuGlyVal 448

Qy       1535 ACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCT 1594
                    |||                  ::::::   |||   |||   ||||||
Db        449 LeuAsnLeuAlaIleValValProGlnMetIleValSerValLeuAlaGlyProPheAsp 468

Qy       1595 GGA-------------------GCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGC 1633
                                         |||      |||      :::   |||
Db        469 SerLeuPheGlyGlyGlyAsnLeuProAlaPheValValGlyAlaIleSerAlaAlaIle 488

Qy       1634 AGTGGCCTGCTCCCA-----------CCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGT 1681
              ||||||:::|||                 |||   |||:::               ||||||
```

```
Qy  1897 AGCAAAGTAAATGGCGACCAGACCC--------------------------AGGCCTGC 1865
Db   150 -----------Gly-ThrArgProProAlaProSerSerArgThrGlyGlnArg-ThrA 165

Qy  1864 GGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCATAAACAG 1805
Db   165 rgGlnArgPro--------------------------GlyHisProProProGluArgG 176

Qy  1804 GGATGGGGCCACCTGGGACA--------------------------------GCAGGAA 1778
Db   176 lySerGlyProArgGlyThrArgProProAlaProSerSerArgThrGlyGlnArgThrA 196

Qy  1777 GGCACTATCCAGGATGGCGAGGTCCAGGCAGATGCCCCGGCCCG---------GAACCAC 1727
Db   196 rgGlnArgProGlyHisProProProGluArgGlySerGlyProArgGlyThrArgProP 216

Qy  1726 CCTGGCCT-----------------------------CGGTGGGCTCACCCACCACCA- 1698
Db   216 roAlaProSerSerArgThrGlyGlnArgThrArgGlnArgProGlyHisProProProG 236

Qy  1697 ----------------------------------------------------------C 1697
Db   236 luArgGlySerGlyProArgGlyThrArgProProAlaProSerSerArgThrGlyGlnA 256

Qy  1696 ACGTACGGAGACATCACAGGCAGAGGCCCCGCAGAGCGCGGGTGGAGGTGGGAGCAGGCC 1637
Db   256 rgThrArgGlnArg-ProGlyHisProProProGluArg-GlySerGlyProArgGlyTh 275

Qy  1636 ACTGCCTCCAGCACCCACGTGTCCATTAGGGAAG-----------------GGAGCTCC 1595
Db   275 rArgProProAlaProSerSerArgThrGlyGlnArgThrArgGlnArgProGlyHisPr 295

Qy  1594 A------------GGCTTAGGGCCT---GGCAGGAAGCTGGTCATCAGGCTGTCCTCACT 1550
Db   295 oProProGluArgGlySerGlyProArgGlyThrArg----------------------- 307

Qy  1549 GCTAGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGTA 1490
Db   308 -------ProProAlaProSerSerArgThrGlyGlnArgThr-ArgGlnArgProGlyH 325

Qy  1489 GAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGC 1430
Db   325 isProProProGluArgGlySerGlyProArgGlyThrArg------------------- 338

Qy  1429 GGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACCGGCAG------- 1377
Db   339 ----------ProProAlaProSerSerArgThrGlyGlnArgThrArgGlnArgProG 355

Qy  1376 ----------------------------CCACAGGGAAAGCTGCCACACTGGCCA---A 1349
Db   355 lyHisProProProGluArgGlySerGlyProArgGlyThrArgProProAlaProSerS 375

Qy  1348 ATAGACTGCTCGAGTGCCGAATCGCTGCACCAGCCGGTCCATGACCAGAGAGAAGACCAG 1289
Db   375 erArgThrGlyGlnArgThrArgGlnArgProGlyHisProProProGluArgGlySerG 395

Qy  1288 GGAGATGGCGCACTGCAGGAACAGCCCCAGGCTGCCCATCCGAACGCCTTCATCATAGTG 1229
Db   395 ly---------ProArgGlyThrArgProProAlaProSerSerArg------------- 407

Qy  1228 TCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTACAGCCCCTCGCCCAC 1169
Db   408 --ThrGlyGlnArgThrArg------------GlnArgProGlyHisProProProGluA 423

Qy  1168 GAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGCTGCACAGCTCAGCCAC 1109
Db   423 rgGlySerGlyProArgGlyThrArgProProAlaProSerSerArgThrGlyGlnArgT 443

Qy  1108 GAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACAGCTGGTGCAGCCGGGGAAGCAG 1049
Db   443 hrArgGlnArgProGlyHisProProProGluArgGlySerGly--ProArgGlyThrAr 462

Qy  1048 GGCGCCCAGGTTCCGGAAAGCCAAGCGGGCCCGGCATGGACAGCAGTGGGGCGACAAGGA 989
Db   462 gProPro-------------AlaProSerSerArgThrGlyGlnArgThrArgGlnArgPr 478

Qy   988 GGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCTCCTCAGCCACCAG 929
Db   478 oGlyHisProProProGluArgGlySerGlyProArgGlyThrArgProProAlaProSe 498

Qy   928 CAGTGTGGCTGCTACGCAGGTGAGGAAGATG--------------------AGGGTGAG 890
Db   498 rSerArgThrGlyGlnArgThrArgGlnArgProGlyHisProProProGluArgGlySe 518

Qy   889 CAGGCCAAAGAGGCACTCCTCCTGGTGCCCAGGT------------------------A 854
Db   518 r-GlyProArgGlyThrArgProProAlaProSerSerArgThrGlyGlnArgThrArgG 538

Qy   853 GGGGGCCAGGGCACTGGTGTCCCAGTCAATGGCAGGCA------GGAGGTAGCCCAGGCA 800
Db   538 lnArgProGlyHisProProProGluArgGlySerGlyProArgGlyThrArgProProA 558

Qy   799 GCCCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGCGACAGTGGTCCGGGTC 740
Db   558 laProSerSerArgThrGlyGlnArgThrArgGlnArgProGlyHisProProProGluA 578

Qy   739 CCGGAAGAGGTCAGAGA---GCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAA 683
Db   578 rgGlySerGlyProArgGlyThrArgProProAlaProSerSerArgThrGlyGlnArgT 598

Qy   682 GTCCAGCAGCCCCACGCCCAGGATGAGCAGTGCCAGCTCCAGGGGCCTGGGATCCGGGCA 623
Db   598 hr------ArgGlnArgProGlyHisProProProGluArgGlySerGlyProArgGlyT 616

Qy   622 CAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCAA 563
Db   616 hr-----------------------------------ArgProProAlaProSerSerA 624

Qy   562 GGACAGTGCCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGC 503
Db   624 rgThrGlyGlnArgThrArgGlnArgProGlyHisProProProGluArgGly------- 641

Qy   502 TGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCATGGT 443
Db   642 ----------------------SerGlyProArgGlyThrArgProProAlaProSerS 654

Qy   442 CATGAACT--------------------------------------------------- 435
Db   654 erArgThrGlyGlnArgThrArgGlnArgProGlyHisProProProGluArgGlySerG 674

Qy   434 --------TCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGC 383
Db   674 lyProArgGlyThrArgProProAlaProSerSerArgThrGlyGlnArgThrArgGlnA 694

Qy   382 GGCCAAACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCA------------------ 342
Db   694 rgProGlyHisProProProGluArgGlySerGly---ProArgGlyThrArgProProA 713

Qy   341 --------GCAAGAGCTGGGCTTTCCGGTGCCGCAGCAGGCGGCTCACCCACAGCCTCTG 290
Db   713 laProSerSerArgThrGlyGlnArg-ThrArgGlnArgProGlyHisProProProGlu 732

Qy   289 GACCATAGTGGGCCAGGCGGG---------TAGGGCTCAGGGGGCCGTTCAGGC------ 245
Db   733 ArgGlySerGlyProArgGlyThrArgProProAlaProSerSerArgThrGlyGlnArg 752

Qy   244 ------------------------ACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCAGAA 209
Db   753 ThrArgGlnArgProGlyHisProProProGlu---------ArgGlySerGlyProArg 769

Qy   208 GCTGCGGCCTCTCCT---CCTTGCTGCCGC------CAACTGCCTAGGAATCAGCCAGGC 158
Db   770 GlyThrArgProProAlaProSerSerArgThrGlyGlnArgThrArgGlnArgProGly 789

Qy   157 GCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCA---GCTTCTCAGCCCATGCTCAACACC 101
```

```
Qy      1529 GGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCT 1588
             |||          |||          ::::::   |||   |||   ||||||
Db       447 GlyValLeuAsnLeuAlaIleValValProGlnMetIleValSerValLeuAlaGlyPro 466

Qy      1589 AAGCCTGGA--------------------GCTCCCTTCCCTAATGGACACGTGGGTGCT 1627
                                          |||    |||   :::   |||
Db       467 PheAspSerLeuPheGlyGlyGlyAsnLeuProAlaPheValValGlyAlaIleSerAla 486

Qy      1628 GGAGGCAGTGGCCTGCTCCCA------------CCTCCACCCGCGCTCTGCGGGGCCTCT 1675
                   ||||||:::|||                |||   |||:::        ||||||
Db       487 AlaIleSerGlyValLeuAlaIleValLeuLeuProLysProSerLysAspAlaAlaSer 506

Qy      1676 GCCTGTGATGTCTCC 1690
                      :::|||
Db       507 LysLeuSerLeuSer 511
```

```
RESULT 18
Q8UZE1
ID   Q8UZE1      PRELIMINARY;      PRT;   890 AA.
AC   Q8UZE1;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   LF3.
OS   cercopithicine herpesvirus 15.
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae; Lymphocryptovirus.
OX   NCBI_TaxID=104228;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LCL8664;
RX   MEDLINE=97048062; PubMed=8892903;
RA   Franken M., Devergne O., Rosenzweig M., Annis B., Kieff E., Wang F.;
RT   "Comparative analysis identifies conserved tumor necrosis factor
RT   receptor-associated factor 3 binding sites in the human and simian
RT   Epstein-Barr virus oncogene LMP1.";
RL   J. Virol. 70:7819-7826(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LCL8664;
RX   MEDLINE=99412410; PubMed=10482645;
RA   Rivailler P., Quink C., Wang F.;
RT   "Strong selective pressure for evolution of an Epstein-Barr virus
RT   LMP2B homologue in the rhesus lymphocryptovirus.";
RL   J. Virol. 73:8867-8872(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LCL8664;
RX   MEDLINE=20304984; PubMed=10846073;
RA   Jiang H., Cho Y.G., Wang F.;
RT   "Structural, functional, and genetic comparisons of Epstein-Barr virus
RT   nuclear antigen 3A, 3B, and 3C homologues encoded by the rhesus
RT   lymphocryptovirus.";
RL   J. Virol. 74:5921-5932(2000).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LCL8664;
RX   MEDLINE=20440633; PubMed=10970361;
RA   Rao P., Jiang H., Wang F.;
RT   "Cloning of the rhesus lymphocryptovirus viral capsid antigen and
RT   Epstein-Barr virus-encoded small RNA homologues and use in diagnosis
RT   of acute and persistent infections.";
RL   J. Clin. Microbiol. 38:3219-3225(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LCL8664;
RX   MEDLINE=21602573; PubMed=11739708;
RA   Rivailler P., Jiang H., Cho Y.-G., Quink C., Wang F.;
RT   "Complete Nucleotide Sequence of the Rhesus Lymphocryptovirus: Genetic
RT   Validation for an Epstein-Barr Virus Animal Model.";
RL   J. Virol. 76:421-426(2002).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LCL8664;
RA   Moghaddam A., Koch J., Annis B., Wang F.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LCL8664;
RA   Moghaddam A., Annis B., Wang F.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
RN   [8]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LCL8664;
RA   Rivailler P., Quink C., Wang F.;
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
RN   [9]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LCL8664;
RA   Jiang H., Wang F.;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [10]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LCL8664;
RA   Rao P.V., Jiang H., Wang F.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [11]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LCL8664;
RA   Rivailler P., Jiang H., Cho Y.-G., Quink C., Wang F.;
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY037858; AAK95470.1; -.
SQ   SEQUENCE   890 AA;  95097 MW;  149077C89D3B5574 CRC64;

Alignment Scores:
Pred. No.:               2.98e-14     Length:         890
Score:                   347.00       Matches:        246
Percent Similarity:      32.15%       Conservative:   52
Best Local Similarity:   26.54%       Mismatches:     361
Query Match:             5.57%        Indels:         268
DB:                      12           Gaps:           41

US-09-759-143-110 (1-3410) x Q8UZE1 (1-890)
```

```
Qy      2341 CCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCTCAACTAGGA 2282
             ||||||:::                             :::|||:::   |||   |||
Db        17 ProGlyGluProAsnGlnArgThrGlnGlyHisProAlaProGlyProSerSerArgThr 36

Qy      2281 GGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCAGAGTCCCCGCATTCCAGTGCA 2222
             |||                           |||:::||||||   |||
Db        37 GlyGlnArgThrArgGlnArgProGlyHisProProProGlu------------ArgGly 52

Qy      2221 TGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGCCTCCAG 2162
                      |||   |||                  |||      |||   :    :::
Db        53 SerGlyProArgGly------------------ThrArgProProAlaProSerSerArg 66

Qy      2161 TCAGGCAGC------------------------------CCTAGAGACTGGGGAGAGAGG 2132
             :::|||                                 |||
Db        67 ThrGlyGlnArgThrArgGlnArgProGlyHisProProProGluArgGlySerGlyPro 86

Qy      2131 AGAGGGACGCCCCAGCCCCCAGCTGTGCAGCTACGCACCTCA------GCAGCACAGGGT 2078
             |||||||||   :::|||||||||         ||||||             |||
Db        87 ArgGlyThr---ArgProProAlaProSerSerArgThrGlyGlnArgThrArgGlnArg 105

Qy      2077 GGCAGCAGAGAGCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCC 2018
                      |||         ||||||         |||      |||   |||
Db       106 ProGlyHisProProProGluArgGlySerGlyPro------ArgGlyThrArgProPro 123

Qy      2017 CATGGGGCTAACAGGAGCGGGGAGCTGGGACCCAGTGAGGCAGGCCCTCCACCCCAATGT 1958
                :::::|||:::|||:::                             |||   ||||||
Db       124 AlaProSerSerArgThrGlyGlnArgThrArgGlnArgProGlyHisProProProGlu 143

Qy      1957 GCTGGAAGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAAATACTACCTGTGT 1898
                ||||||      |||
Db       144 ArgGlySerGlyProArg------------------------------------------ 149
```

```
Qy       1529 GGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCT 1588
              |||         |||         ::::::  |||   |||   ||||||
Db        447 GlyValLeuAsnLeuAlaIleValValProGlnMetIleValSerValLeuAlaGlyPro 466

Qy       1589 AAGCCTGGA--------------------GCTCCCTTCCCTAATGGACACGTGGGTGCT 1627
                                           |||     |||    :::   |||
Db        467 PheAspSerLeuPheGlyGlyGlyAsnLeuProAlaPheValValGlyAlaIleSerAla 486

Qy       1628 GGAGGCAGTGGCCTGCTCCCA-----------CCTCCACCCGCGCTCTGCGGGGCCTCT 1675
                    |||||||:::|||         |||   |||:::         ||||||
Db        487 AlaIleSerGlyValLeuAlaIleValLeuLeuProLysProSerLysAspAlaAlaSer 506

Qy       1676 GCCTGTGATGTCTCC 1690
                       :::|||
Db        507 LysLeuSerLeuSer 511

RESULT 17
Q9FNR6
ID   Q9FNR6      PRELIMINARY;      PRT;   515 AA.
AC   Q9FNR6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Sucrose/proton symporter.
GN   SUT2.
OS   Daucus carota (Carrot).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
OX   NCBI_TaxID=4039;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. NANTAISE;
RA   Sturm A.;
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. NANTAISE;
RA   Shakya R.;
RL   Thesis (2000), Department of Department of Botany,
RL   University of Basel, Basel, Switzerland.
DR   EMBL; AJ303199; CAC19689.1; -.
DR   InterPro; IPR003662; sub_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   TIGRFAMs; TIGR01301; GPH_sucrose; 1.
SQ   SEQUENCE   515 AA;  54425 MW;  9580978C726523DC CRC64;

Alignment Scores:
Pred. No.:              2.53e-14    Length:         515
Score:                  347.50      Matches:        131
Percent Similarity:     41.78%      Conservative:   80
Best Local Similarity:  25.94%      Mismatches:     215
Query Match:            5.41%       Indels:         79
DB:                     10          Gaps:           17

US-09-759-143-110 (1-3410) x Q9FNR6 (1-515)

Qy        332 CAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATC 391
              :::|||:::||||||       :::       |||::::::        |||      :::
Db         34 LysLeuValLeuValAlaAlaIleAlaAlaGlyValGlnPheGlyTrpAlaLeuGlnLeu 53

Qy        392 ACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTG 451
              :::   :::   |||   :::      :::|||:::      |||:::            :::
Db         54 SerLeuLeuThrProTyrValGlnLeuLeuGlyIleProHisLysTrpAlaAlaTyrIle 73

Qy        452 CTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGAC 511
                             ||||||:::   |||::::::      |||:::::::|||      ||||||
Db         74 TrpLeuCysGlyProIleSerGlyMetLeuValGlnProIleValGlyTyrTyrSerAsp 93

Qy        512 CACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATC 571
              |||   :::         :::||||||||||||||||||||   :::   :::      :::
Db         94 HisCysGlnSerSerPheGlyArgArgArgProPheIleAlaSerGlyAlaGlyCysVal 113

Qy        572 CTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTG----------- 619
                 :::|||:::  ||||||         |||:::  |||   :::
Db        114 AlaIleSerValIleLeuIle---------GlyPheAlaAlaAspIleSerTyrLysAla 130

Qy        620 ---------CTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTQATCCTGGGCGTG 670
                        :::         :::|||   :::   :::   ::::::|||
Db        131 GlyAspAspMetSerLysThrLeuLysProArgAlaValThrValPheValIleGlyPhe 150

Qy        671 GGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGAC 730
                 :::||||||         :::         |||      |||||||||:::|||
Db        151 TrpIleLeuAspValAlaAsnAsnMetLeuGlnGlyProCysArgAlaLeuLeuAlaAsp 170

Qy        731 CTCTTCCGG---GACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATC 787
              |||      |||            |||   |||   :::   |||:::|||      :::
Db        171 LeuCysSerGlyAspThrArgArgMetArgSerAlaAsnAlaPheTyrSerPhePheMet 190

Qy        788 AGTCTTGGGGGCTGCCTGGGCTAC----------------------------CTCCTG 817
              ::::::|||      |||||||||                           |||
Db        191 AlaValGlyAsnIleLeuGlyTyrAlaAlaGlySerTyrAsnAsnLeuTyrLysLeuPhe 210

Qy        818 CCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGC 877
              |||         |||   |||         |||                  :::   |||
Db        211 ProPheSer------LysThrHisAlaCysAspLeuTyrCysAlaAsnLeuLysSerCys 228

Qy        878 CTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCT 937
                       :::   |||:::   :::         |||   |||      :::|||
Db        229 PheIleIleSerIleAlaLeuLeuIleIleIleThrValValAlaLeuSerValValArg 248

Qy        938 GAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCG 997
              |||   :::      ||||||   :::   |||:::                  |||
Db        249 GluAsnSer------GlyProProAspAspAlaAspAlaAlaGluGluPro--------- 263

Qy        998 CCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCC 1057
                          |||      ::::::                  |||   ||||||
Db        264 -----------ProSerSerGlyLysIleProVal------PheGlyGluLeuLeuGly 277

Qy       1058 CGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAG 1117
                 |||   |||               ||||||   :::   |||   :::
Db        278 AlaLeuLysAspLeu-----------ProArgProMetLeuLeuLeuIleValThr 293

Qy       1118 CTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAG 1177
                    :::|||:::|||         |||   ||||||   ||||||::::::|||
Db        294 CysLeuAsnTrpIleAlaTrpPheProPheIleLeuPheAspThrAspTrpMetGlyArg 313

Qy       1178 GGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGAT 1237
                 :::|||   |||         |||   |||                  :::   ||||||
Db        314 GluIleTyrGlyGlyThr------AlaGlyLysGly-----------LysLeuTyrAsp 327

Qy       1238 GAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTC 1297
              :::|||||||||   |||:::|||||||||   |||         :::      :::
Db        328 GlnGlyValArgAlaGlyAlaLeuGlyLeuLeuLeuAsnSerValValLeuGlyLeuThr 347

Qy       1298 TCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGT 1357
              |||:::   ::::::   ||||||:::         |||         :::   |||
Db        348 SerIleAlaValGluTyrLeuValArgGlyValGlyGlyValLysIle---LeuTrpGly 366

Qy       1358 GTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTG-------------------- 1396
                    |||   :::|||   |||
Db        367 PheValAsnPheIleLeuAlaIleGlyLeuValMetThrValValValSerLysValAla 386

Qy       1397 -----------TCCCACAGTGTG----------------------GCCGTGGTGACA 1420
                         ||||||                          |||   |||
Db        387 GlnHisGlnArgGluHisSerAlaAsnGlyGlnLeuLeuProProSerAlaGlyValLys 406

Qy       1421 GCTTCAGCC-----------GCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTG 1468
              |||   |||              ::::::   |||         |||            :::
Db        407 AlaGlyAlaLeuSerLeuPheSerIleLeuGlyIleProLeuSerIleThrTyrSerIle 426

Qy       1469 CCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGA 1528
              |||:::   |||||||||:::|||
Db        427 ProPheAlaLeuAlaSerIleTyrSerSerGlySerGlyAlaGlyGlnGlyLeuSerLeu 446
```

```
                 |||          :::    :::    :::                |||
Db         405 PheSerIleLeuGlyIleProLeuAlaIleThrTyrSerValProTyrAlaLeuIleSer 424

Qy        1721 GCCAGGGTGGTTCCG---GGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCC 1777
                  |||:::           |||    |||       |||    :::    :::|||:::    |||
Db         425 ThrArgIleGluSerLeuGlyLeuGlyGlnGlyLeuSerMetGlyValLeuAsnLeuAla 444

Qy        1778 TTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGCTCCATTGTCCAGCTC------ 1831
                 ::::::   ||||||         ||||||       |||         ||||||
Db         445 IleValIleProGlnValIleValSerLeuGlySerGlyProTrpAspGlnLeuPheGly 464

Qy        1832 ---------AGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTC 1882
                           |||   |||:::            |||:::|||         |||||||||
Db         465 GlyGlyAsnSerProSerLeuAlaValAlaAlaValAlaAlaPheAlaSerGlyLeuVal 484

Qy        1883 GCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGC 1921
               ||||||            :::         |||||||||
Db         485 AlaIleLeuAlaIleProArgSerSerAlaAspLysSer 497

RESULT 16
O65803
ID   O65803      PRELIMINARY;      PRT;   515 AA.
AC   O65803;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Sucrose/H+ symporter.
GN   SUT2.
OS   Daucus carota (Carrot).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
OX   NCBI_TaxID=4039;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. NAMTAISE; TISSUE=ROOT;
RX   MEDLINE=99063785; PubMed=9847123;
RA   Shakya R., Sturm A.;
RT   "Characterization of source- and sink-specific sucrose/H+ symporters
RT   from carrot.";
RL   Plant Physiol. 118:1473-1480(1998).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR   EMBL; Y16768; CAA76369.1; -.
DR   InterPro; IPR003662; sub_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   TIGRFAMs; TIGR01301; GPH_sucrose; 1.
KW   Transmembrane.
SQ   SEQUENCE   515 AA;  54424 MW;  897B90657C9E243C CRC64;

Alignment Scores:
Pred. No.:               2.53e-14       Length:         515
Score:                   347.50         Matches:        131
Percent Similarity:      41.78%         Conservative:   80
Best Local Similarity:   25.94%         Mismatches:     215
Query Match:             5.41%          Indels:         79
DB:                      10             Gaps:           17

US-09-759-143-110 (1-3410) x O65803 (1-515)

Qy         332 CAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATC 391
               :::|||:::||||||      :::         |||::::::           |||       :::
Db          34 LysLeuValLeuValAlaAlaIleAlaAlaGlyValGlnPheGlyTrpAlaLeuGlnLeu 53

Qy         392 ACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTG 451
               :::    :::    |||    :::         :::|||:::         |||:::         :::
Db          54 SerLeuLeuThrProTyrValGlnLeuLeuGlyIleProHisLysTrpAlaAlaTyrIle 73

Qy         452 CTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGAC 511
                           |||||||:::    |||::::::        |||:::::::|||        ||||||
Db          74 TrpLeuCysGlyProIleSerGlyMetLeuValGlnProIleValGlyTyrTyrSerAsp 93

Qy         512 CACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATC 571
               |||    :::         :::||||||||||||||||||||||       :::    :::       :::
Db          94 HisCysGlnSerSerPheGlyArgArgArgProPheIleAlaSerGlyAlaGlyCysVal 113

Qy         572 CTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTG------------ 619
                  :::|||:::    ||||||            |||:::    |||    :::
Db         114 AlaIleSerValIleLeuIle---------GlyPheAlaAlaAspIleSerTyrLysAla 130

Qy         620 ---------CTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTG 670
                           :::          :::|||       :::    :::    ::::::|||
Db         131 GlyAspAspMetSerLysThrLeuLysProArgAlaValThrValPheValIleGlyPhe 150

Qy         671 GGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGAC 730
                  :::||||||          :::          |||       |||||||||:::|||
Db         151 TrpIleLeuAspValAlaAsnAsnMetLeuGlnGlyProCysArgAlaLeuLeuAlaAsp 170

Qy         731 CTCTTCCGG---GACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATC 787
               |||       |||       |||    |||    :::    |||:::|||    :::
Db         171 LeuCysSerGlyAspThrArgArgMetArgSerAlaAsnAlaPheTyrSerPhePheMet 190

Qy         788 AGTCTTGGGGGCTGCCTGGGCTAC----------------------------CTCCTG 817
               ::::::|||       |||||||||                              |||
Db         191 AlaValGlyAsnIleLeuGlyTyrAlaAlaGlySerTyrAsnAsnLeuTyrLysLeuPhe 210

Qy         818 CCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGC 877
               |||             |||    |||          |||                :::    |||
Db         211 ProPheSer------LysThrHisAlaCysAspLeuTyrCysAlaAsnLeuLysSerCys 228

Qy         878 CTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCT 937
                           :::    |||:::    :::          |||    |||       :::|||
Db         229 PheIleIleSerIleAlaLeuLeuIleIleIleThrValValAlaLeuSerValValArg 248

Qy         938 GAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCG 997
               |||    :::          ||||||    :::    |||:::                |||
Db         249 GluAsnSer------GlyProProAspAspAlaAspAlaAlaGluGluPro--------- 263

Qy         998 CCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCC 1057
                            |||       ::::::                    |||    ||||||
Db         264 ------------ProSerSerGlyLysIleProVal------PheGlyGluLeuLeuGly 277

Qy        1058 CGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAG 1117
                  |||    |||          ||||||    :::       |||    :::
Db         278 AlaLeuLysAspLeu------------ProArgProMetLeuLeuLeuLeuIleValThr 293

Qy        1118 CTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAG 1177
                  :::|||:::|||          |||    ||||||       ||||||::::|||
Db         294 CysLeuAsnTrpIleAlaTrpPheProPheIleLeuPheAspThrAspTrpMetGlyArg 313

Qy        1178 GGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGAT 1237
                  :::|||    |||          |||       |||                :::    ||||||
Db         314 GluIleTyrGlyGlyThr------AlaGlyGlnGly-----------LysLeuTyrAsp 327

Qy        1238 GAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTC 1297
               :::|||||||||    |||:::|||||||||    |||                :::       :::
Db         328 GlnGlyValArgAlaGlyAlaLeuGlyLeuLeuLeuAsnSerValValLeuGlyLeuThr 347

Qy        1298 TCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGT 1357
               |||:::    ::::::    ||||||:::       |||                :::    |||
Db         348 SerIleAlaValGluTyrLeuValArgGlyValGlyGlyValLysIle---LeuTrpGly 366

Qy        1358 GTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTG-------------------- 1396
                           |||    :::|||    |||
Db         367 PheValAsnPheIleLeuAlaIleGlyLeuValMetThrValValValSerLysValAla 386

Qy        1397 -----------TCCCACAGTGTG----------------------GCCGTGGTGACA 1420
                                ||||||                            |||    |||
Db         387 GlnHisGlnArgGluHisSerAlaAsnGlyGlnLeuLeuProProSerAlaGlyValLys 406

Qy        1421 GCTTCAGCC-----------GCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTG 1468
               |||    |||                   ::::::    |||          |||                :::
Db         407 AlaGlyAlaLeuSerLeuPheSerIleLeuGlyIleProLeuSerIleThrTyrSerIle 426

Qy        1469 CCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGA 1528
               |||:::    |||||||||:::|||
Db         427 ProPheAlaLeuAlaSerIleTyrSerSerGlySerGlyAlaGlyGlnGlyLeuSerLeu 446
```

```
Db        747 alArg 748

RESULT 15
Q9SQK6
ID   Q9SQK6      PRELIMINARY;      PRT;   501 AA.
AC   Q9SQK6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Putative sucrose transporter.
GN   VVSUC11.
OS   Vitis vinifera (Grape).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Vitaceae;
OC   Vitis.
OX   NCBI_TaxID=29760;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SHIRAZ; TISSUE=FRUIT;
RA   Davies C., Wolf T., Robinson S.P.;
RT   "Three putative sucrose transporters are differentially expressed in
RT   grapevine tissues.";
RL   Plant Sci. 147:93-100(1999).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR   EMBL; AF021808; AAF08329.1; -.
DR   InterPro; IPR003662; sub_transporter.
DR   Pfam; PF00083; sugar_tr; 1.
DR   TIGRFAMs; TIGR01301; GPH_sucrose; 1.
KW   Transmembrane.
SQ   SEQUENCE   501 AA;  53938 MW;  4D0D4DE2EF2F4BA8 CRC64;

Alignment Scores:
Pred. No.:              2.17e-14      Length:         501
Score:                  348.50        Matches:        164
Percent Similarity:     40.62%        Conservative:   85
Best Local Similarity:  26.75%        Mismatches:     205
Query Match:            5.43%         Indels:         159
DB:                     10            Gaps:           25

US-09-759-143-110 (1-3410) x Q9SQK6 (1-501)

Qy        191 AGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTCTGGAGTGCCTGA 250
              |||:::|||||||||
Db          8 ArgGlnArgGlyArg--------------------------------------------- 12

Qy        251 ACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGTGAGCCGC 310
                    |||  ||||||         |||         |||:::   :::   |||
Db         13 ------ProArgAlaLeuIleGlyGluProValArg---ProArgValProLeuArgArg 29

Qy        311 CTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAG 370
              |||||||||                       |||     :::     |||::::::
Db         30 LeuLeuArg-----------------------ValAlaSerValAlaCysGlyIleGln 41

Qy        371 GTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAG 430
                       |||   ::::::  :::  |||   :::  |||:::|||:::
Db         42 PheGlyTrpAlaLeuGlnLeuSerLeuLeuThrProTyrValGlnGluLeuGlyIlePro 61

Qy        431 GAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCG 490
                 :::  :::::::::             ||||||:::  ||||||:::        |||
Db         62 HisAlaTrpSerSerIleIleTrpLeuCysGlyProLeuSerGlyLeuLeuValGlnPro 81

Qy        491 CTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATC 550
              |||:::|||         ||||||         |||:::||||||||||||||||||||||
Db         82 LeuValGlyHisLeuSerAspArgCysAsnSerArgPheGlyArgArgArgProPheIle 101

Qy        551 TGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTA 610
                 |||   :::   |||::::::::::::   :::|||        :::        :::
Db        102 ValAlaGlyAlaThrSerIleValValAlaValLeuIleIleGlyPheSerAlaAspIle 121

Qy        611 GCAGGGCTGCTG------TGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTG 664
                 |||||||||         |||  ||||||        :::|||        ::::::
Db        122 GlyGlyLeuLeuGlyAspGlyAlaAspArgArgProArgAlaValAlaThrPheValVal 141

Qy        665 GGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTC 724
              |||      |||||||||          |||          |||      |||||||||
Db        142 GlyPheTrpLeuLeuAspValAlaAsnAsnValThrGlnGlyProCysArgAlaLeuLeu 161

Qy        725 TCTGACCTCTTCCGGGACCCGGACCAC------TGTCGCCAGGCCTACTCTGTCTATGCC 778
              :::||||||      :::   ||||||       |||    |||    :::    ::::::
Db        162 AlaAspLeu---ThrGluLysAspHisArgArgThrArgValAlaAsnAlaTyrPheSer 180

Qy        779 TTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC--------------- 823
                 |||::::::|||      ||||||:::          :::
Db        181 LeuPheIleAlaValGlyAsnValLeuGlyPheAlaThrGlySerTyrSerGlyTrpPhe 200

Qy        824 ---ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTC 880
                 |||   |||   ||||||        :::                        :::|||
Db        201 ArgIlePheTrpPheThrSerThrSerSer-----------CysAsnAlaAspCysAla 216

Qy        881 -----------TTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTG 928
                          ||||||   :::||||||         :::|||   |||
Db        217 AsnLeuLysSerAlaPheLeuLeuAspIleIlePhe---------IleAlaIleThrThr 233

Qy        929 CTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCC 988
                 ::::::      ||||||                       ||||||:::
Db        234 TyrIleSerIleThrAlaAla--------------GlnGluLeuProLeuSerSerSer 248

Qy        989 TCCTTGTCGCCCCACTGCTGT--------------CCATGCCGGGCCCGCTTGGCTTTC 1033
              |||   |||   |||                                |||:::   ||||||
Db        249 SerArgSerThrHisIleSerGluGluMetAlaGluSerThrHisAlaGlnGluAlaPhe 268

Qy       1034 CGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACC 1093
                                           |||   :::|||                 |||
Db        269 --------------------------LeuTrpGluLeuPhe-----------GlyThr 275

Qy       1094 CTGCGCCGG----------------------CTCTTCGTGGCTGAGCTGTGCAGCTGG 1129
              ||||||                         |||||||||       |||   :::|||
Db        276 LeuArgTyrLeuSerGlySerIleTrpIleIleLeuPheValThrAlaLeu---ThrTrp 294

Qy       1130 ATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAG 1189
              :::   |||:::   |||   ||||||   ||||||::::::|||         :::|||
Db        295 IleGlyLeuLeuProPheLeuLeuPheAspThrAspTrpMetGlyArgGluIleTyrGly 314

Qy       1190 GGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGG 1249
              |||   |||                        ::::::|||:::   |||||||||
Db        315 GlyLysPro-----------------AsnGluGlyGlnAsnTyrAsnThrGlyValArg 328

Qy       1250 ATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATG 1309
              ||||||:::||||||||||   |||         :::     :::   |||::::::|||
Db        329 MetGlyAlaLeuGlyLeuMetLeuAsnSerValValLeuGlyIleThrSerValLeuMet 348

Qy       1310 GACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTC 1369
              ::::::|||   :::::::::|||         |||:::
Db        349 GluLysLeuCysArgLysTrpGlyAlaGlyPheValTrp-------------------- 361

Qy       1370 CCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCC 1429
                                                                   |||
Db        362 ------------------------------------------------GlyLeuSerAsn 365

Qy       1430 GCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTC 1489
                 |||         |||   :::   ||||||                 |||   :::
Db        366 IleLeuMetSerLeuCysPheLeuLeuMetLeuIleLeuSer---------AlaValVal 382

Qy       1490 TACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGC 1549
                 |||   :::
Db        383 LysHisMetAsp------------------------------------------------ 386

Qy       1550 AGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCT 1609
                                       ||||||         |||                |||
Db        387 -----------------------PheLeuGlyHisAspLeuPro------------Pro 394

Qy       1610 AATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGG 1669
              :::|||   |||   |||                              ||||||
Db        395 SerGlyValValIleAla-----------------------------AlaLeuIleVal 404

Qy       1670 GCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCC---------ACCGAG 1720
```

```
Qy       1868 CTGCGGCAGACACCATATAGGCAGTGACAGACTGGCTGAGCTGGACAATGGAGCCCATAA 1809
                :::     |||            :::   |||   ||||||        |||
Db        186 PheGlnGlyProProGlyGluProGlyGluProGlyAlaAlaGlyAlaMetGlyPro--- 204

Qy       1808 ACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCAC------------------------ 1773
                    ||||||||||||       |||::::::
Db        205 ------ArgGlyProProGlyProProGlyLysAsnGlyGluAspGlyGluAlaGlyLys 222

Qy       1772 ------TATCCAGGATGGCGAGGTCCAGGCAGATGCCCCGGCCCGGAA------------ 1731
                    ||||||   |||||||||         |||||||||:::
Db        223 ProGlyArgProGlyGluArgGlyPro---------ProGlyProGlnGlyAlaArgGly 239

Qy       1730 ---------CCACCCTGGCCTCGGTGGGCTCACCCACCACCACACGTAC----------- 1691
                           ||||||   |||:::      |||        :::
Db        240 LeuProGlyThr-AlaGlyLeuProGlyMetLysGlyHisArgGlyPheAsnGlyLeuAs 259

Qy       1690 ---------GGAGACATCACAGGCAGAGGCCCCGCAGAGCGCGGGTGGAGGT-------- 1648
                        ||||||            ||| |||:::        |||
Db        259 pGlyAlaLysGlyAspThrGlyProAlaGly-ProLysGlyGluProGlyAsnProGlyG 279

Qy       1647 ----------------------GGGAGCAGGCCACTGCCT---------------CCAG 1626
                                    |||   |||   ||||||              |||
Db        279 luAsnGlyAlaProGlyGlnValGlyProArgGlyLeuProGlyGluArgGlyArgProG 299

Qy       1625 CACCCACGTGTCCATTAGGG------------AAGGGAGCTCCAGGCTTAGGGCCTGGCA 1578
                |||:::    |||    |||                  |||   ||||||
Db        299 lyProSerGlyProAlaGlyAlaArgGlyAsnAspGlyThrProGlyAlaAla------- 316

Qy       1577 GGAAGCTGGTCATCAGGCTGTCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGTATTTGG 1518
                                                  ||||||    |||   |||
Db        317 -------------------------------GlyProPro-GlyProThrGlyPro--- 324

Qy       1517 GCAGGAACACCTGCTTCTCCCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCA 1458
                |||   |||       |||||||||   |||   |||          ||||||
Db        325 ThrGlyProProGlyPheProGlyGlyValGly---ProLysGlyAspAlaGlyProGln 343

Qy       1457 GGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGGG 1398
              |||   |||                                |||
Db        344 GlySerArg-----------------------------GlyPro--------------- 348

Qy       1397 ACAGGCATGTGGCACCGGCAGCCACAGGGAAAGCTG------CCACACTGGCCAAATAGA 1344
                                 |||||||||         |||      |||    :::
Db        349 --------------AspGlyProGlnGlyAlaArgGlyGluProGlyAlaProGlyGln 363

Qy       1343 CTGCTCGAGTGCCGAATCGCTGCACCAGCCGGT-----------CCATGACCAGAGAGA 1296
                                 |||   |||||||||         |||      :::
Db        364 -----------------AlaGlyProAlaGlySerAlaGlyAsnProGlyThrAspGly 377

Qy       1295 AGACCAGGGAGATGGCGCACTGCAGGAACA--------------GCCCCAGGCTGCCCA 1251
              :::||||||              |||                  |||||||||   |||
Db        378 GlnProGlyAlaLysGlyAlaThrGlyAlaProGlyIleAlaGlyAlaProGlyPhePro 397

Qy       1250 TCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGC 1191
                |||              ||||||:::
Db        398 GlyAlaArg---------GlyAlaProGlyProGln--------------------Gly 407

Qy       1190 CCTGGTACAGCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCA 1131
              ||||||   :::|||   |||:::
Db        408 ProGlyGlySerProGlyProLysGlyAsn---------------------------- 417

Qy       1130 TCCAGCTGCACAGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACA 1071
                           :::|||                         ||||||   |||:::
Db        418 ---------AsnGlyGluPro--------------------GlyAlaGlnGlyAsnLys 427

Qy       1070 GCT------GGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCC---GGAAAGCCAAGCGGG 1020
                     ||||||   ||||||:::|||         ||||||   |||   |||
Db        428 GlyGluProGlyAlaLysGlyGluSerGlyProAlaGlySerGlnGlyProProGlyPro 447

Qy       1019 CCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGG 960
              ||||||   :::           :::|||      |||
Db        448 ProGlyGluGluGlyLysArgGlySerArgGlyGluPro-------------------- 460

Qy        959 TGGGGCCCAGCGCTGCCTCCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGA 900

                 ||||||:::    ||||||     |||
Db        461 ---GlyProSerGlyProProGlyPro-------------------------------- 468

Qy        899 TGAGGGTGAGCAGGCCAAAGAGGCACTCCT--------------CCTGGGTGCCCAGGT 855
                         ||||||:::||||||   |||                ||||||   |||
Db        469 ---------AlaGlyGluArgGlyAlaProGlySerArgGlyPheProGlyAlaAspGly 485

Qy        854 AGGGGG--------------CCAGGGCACTGGTGTCCCAGTCAATGGCAGGCAGGAGGT 810
              |||                  ||||||         |||             :::   |||
Db        486 AlaGlyGlyProLysGlyProProGlyGluArgGlyProVal------GlySerAlaGly 503

Qy        809 AGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGCGACAGT 750
                 |||||||||   :::                          |||||||||:::
Db        504 ProLysGlySerProGlyGlu-----------------SerGlyArgProGlyGluPro 517

Qy        749 GGT-------------------------------------CCGGGTCCC---GGAAGA 732
              |||                                      |||||||||   |||:::
Db        518 GlyLeuProGlyAlaLysGlyLeuThrGlySerProGlySerProGlyProAspGlyLys 537

Qy        731 GGTCAGAGAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCC 672
                    ||||||   |||   :::         ||||||         |||
Db        538 ThrGlyProAlaGlyAlaProGlyGlnAspGlyArgProGlyProProGlyProProGly 557

Qy        671 CCACGC----------------------CCAGGATGAGCAGTGCCAGCTCCAGGGGCC 636
              |||                            ||||||         |||      |||
Db        558 AlaArgGlyGlnSerGlyValMetGlyPheProGly---------ProLysGlyAlaAla 574

Qy        635 TGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCA 576
                 :::||||||                              |||||||||
Db        575 GlyGluProGlyLysProGly-------------------------GluArgGlyVal 585

Qy        575 GCAGGATGCCCAAGGACAGTGCCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCC 516
              ||||||   |||      |||
Db        586 AlaGlyProProGlyAlaVal------------------------------------- 592

Qy        515 AGTGGTCACTGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGC 456
                                    ||||||     :::   |||   |||      |||
Db        593 --------------------GlyAlaProGlyLysAspGlyGluAlaGlyAlaGlnGly 605

Qy        455 CCAGCACCATGGTCATGAACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCACAT 396
                                               ||||||   ||||||      |||
Db        606 -----------------------------ProProGlyProAlaGlyProAlaGlyGlu 615

Qy        395 AGGTGATGCCTGCGGCCAAACACACCTCCAGGC---CAAAGGTTAGCAGGTTGACCAGCA 339
              |||         |||      |||||||||   |||   |||   |||   |||
Db        616 ArgGlyGluGlnGlyProAlaGlyProProGlyPheGlnGlyLeuProGlySerProGly 635

Qy        338 AGAGCTGGGCTTTCCGGT---GCCGCAGCAGGCGGCTCACCCACAGCCTCTGGACCATAG 282
                 |||   ||||||                  |||:::|||            |||||| ||
Db        636 AlaProGlyGluSerGlyLysProGlyGluGlnGlyAlaProGlyAspValGlyPro-Se 655

Qy        281 TGGGCCAGGCGGGTAGGGCTCAGGGGGCCGTTCAGGCACTCCA---GAACTGCTTCGTCT 225
              |||||||         ||||||   |||      |||   |||   |||         ::
Db        655 rGlyProAla------GlySerArgGlyGluArgGlyPheProGlyGluArgGlyAlaIl 673

Qy        224 CGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAACTGCCTAGGAATCA 165
              :|||      |||            |||                     |||   |||
Db        673 eGly------ProProGlyProGlnGlyProArgGlyAlaAsnGlyAlaProGlyAsnAs 691

Qy        164 G----------------------CCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGG 129
                                     |||||||||             |||   |||   ||
Db        691 pGlyAlaLysGlyGluAlaGlyAlaProGlyAla-----------ProGlyGlyGlnGl 707

Qy        128 TCCAGCTTCT---CAGCCCATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTGGGGACA 72
              ||||:::      |||   |||                     |||   |||   |||
Db        707 yProSerGlyLeuGlnGlyMet-ProGlyGluArgGlyAlaGlyGlyLeuProGlyAlaL 727

Qy         71 CGTCTCATCACTCAGATCCTGGC---CGAGGCGCGCGGCTGTCACCCGGAGCCAGCGCGT 15
                           |||   |||   :::||||||          :::||||||
Db        727 ysGlyAspArgGlyAspGlnGlyProLysGlyAlaAspGlyAlaProGlyLysAspGlyV 747

Qy         14 GCAGG 10
               |||
```

```
Db        670 eProGlyGluArgGlyValGlnGlyProProGlyProAlaGlyProArgGlyAsnAsnGl 690

Qy        654 AGTGCCA-----------------------------GCTCCAGGGGCCTGGGATCCGGG 625
                 |||                                 |||||||||||    |||||
Db        690 yAlaProGlyAsnAspGlyAlaLysGlyAspThrGlyAlaProGlyAla------ProGl 708

Qy        624 CACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCC 565
              |:::                    :::   :::     |||||||||:::||||||   ||
Db        708 ySerGlnGlyAlaProGlyLeuGlnGlyMetProGlyGluArgGlyAlaAlaGlyLeuPr 728

Qy        564 AAGGACAGTGCCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTG 505
              |         |||:::          |||                 |||   |||
Db        728 oGly------ProLysGlyAspArgGlyAspAlaGlyProLysGlyAlaAspGly----- 744

Qy        504 GCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCACCATG 445
                                                 |||      |||
Db        745 --------------------------SerProGlyLysAspGlyValArgGlyLeu-- 754

Qy        444 GTCATGAACTTCTCCTCTA---CCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATG 388
                     |||   |||:::   ||||||   ||||||              :::
Db        755 -------ThrGlyProIleGlyProProGlyProAlaGlyAlaProGlyAspLysGlyGl 772

Qy        387 CCTGCGGCCAAACACACCT--------------------------CCAGGCCAAAGG-- 357
                 |||:::   |||                              ||||||   |||
Db        772 uAlaGlyProSerGlyProAlaGlyProThrGlyAlaArgGlyAlaProGlyAspArgGl 792

Qy        356 -TTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGT---GCCGCAGCAGGCGGCTCACC 301
                 ||||||   |||      ||||||   :::|||         |||   |||   ||
Db        792 yGluAlaGlyProProGlyProAlaGlyPheAlaGlyProProGlyAlaAspGlyGlnPr 812

Qy        300 CACAGCCTCTGGACCATAGTGGG--------------------CCAGGCGGGTAGGGCT 262
              |   |||   |||         |||                     |||      |||
Db        812 oGlyAlaLysGly-GluProGlyAspThrGlyValLysGlyAspAlaGlyProProGlyP 832

Qy        261 CAGGGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCA---------- 212
                 |||   :::|||   |||        :::         ||||||
Db        832 roAlaGlyProAlaGlyProProGlyProIleGlyAsnValGlyAlaProGlyProLysG 852

Qy        211 -----GAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCAACTGCCTAGGAATCAGCCAGGCG 157
                        ||||||   ||||||   |||         |||         |||
Db        852 lySerArgGlyAlaAlaGlyProPro-GlyAlaThr---------GlyPheProGlyAla 868

Qy        156 CCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGCTTCTCAGCCCATGCTCAACACCTGCT 97
                                  |||      |||           |||         |||
Db        869 AlaGlyArgValGlyProProGlyProSerGlyAsnAlaGlyProProGlyProProGly 888

Qy         96 GCTGTGGGGCACCTCAGT--------------GGGGACACGTCTCATCACTCAGATCCT 52
                 ||||||                          |||:::|||         |||
Db        889 ProValGlyLysGluGlyGlyLysGlyProArgGlyGluThrGlyProAlaGlyArgPro 908

Qy         51 GGCCGA----------------------------GGCGCGCGGCTGTCACCCGGAGCC 22
                 |||                                 |||   :::   |||||||||||
Db        909 GlyGluValGlyProProGlyProProGlyProAlaGlyGluLysGlySerProGlyAla 928
```

RESULT 14
O93251

```
ID   O93251      PRELIMINARY;      PRT;  1445 AA.
AC   O93251;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Alpha 1 type I collagen.
OS   Rana catesbeiana (Bull frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
OX   NCBI_TaxID=8400;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99294154; PubMed=10367734;
RA   Asahina K., Utoh R., Obara M., Yoshizato K.;
RT   "Cell-type specific and thyroid hormone-dependent expression of genes
RT   of a1(I) and a2(I) collagen in intestine duting
RT   amphibianmetamorphosis.";
RL   Matrix Biol. 18:89-103(1999).
DR   EMBL; AB015440; BAA29028.1; -.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR000885; Fib_collagen_C.
DR   InterPro; IPR001007; VWF_C.
DR   Pfam; PF01410; COLFI; 1.
DR   Pfam; PF01391; Collagen; 18.
DR   ProDom; PD000007; Collagen; 3.
DR   ProDom; PD002078; Fib_collagen_C; 1.
DR   SMART; SM00038; COLFI; 1.
DR   SMART; SM00214; VWC; 1.
DR   PROSITE; PS01208; VWFC; 1.
KW   Collagen.
SQ   SEQUENCE   1445 AA;  137252 MW;  F59BB550C9873F04 CRC64;

Alignment Scores:
Pred. No.:              1.06e-14      Length:        1445
Score:                  354.50        Matches:       256
Percent Similarity:     31.75%        Conservative:  51
Best Local Similarity:  26.47%        Mismatches:    307
Query Match:            5.69%         Indels:        353
DB:                     13            Gaps:          58

US-09-759-143-110 (1-3410) x O93251 (1-1445)

Qy       2588 TGTCTCTGTGATGGCAACAGAAGGACCAACAGGCCACATCCTGATAAAAGGTAAGAGGGG 2529
                 |||:::   ||||||                     |||   :::   ||||||
Db         33 CysValValAspGly---------------ArgThrTyrAsnAspLys------------ 43

Qy       2528 GGTGGATCAGCAAAAAGACAGTGCTGTGGGCTGAGGGGACCTGGTTCTTGTGTGTTGCCC 2469
                       |||   :::   |||   :::                    ||||||  |||
Db         44 ---AspValTrpLysProGluAlaCysGlnIle---------------CysVal-CysAs 57

Qy       2468 CTCAGGACTCTTCCCCTACAAATAAGTCATATGTTCAAATCCCATGGAGGAGTGTTTCAT 2409
                 |||            |||||||||   :::      :::   :::|||
Db         57 pGluGlyThrIleLeuCysAspGluValIleCysGluAsp---IleGlyAspCysProAs 76

Qy       2408 CCTAGAAACTCCCATGCAAGAGCTACATTAAACGAAGCTGCAGGTTAAGGGGCTTAGAGA 2349
                 |||   ||||||   |||                       |||   |||
Db         76 nProGluIleProMetGlyGluCys------------CysProValCysGlyGluGlyGl 92

Qy       2348 TGGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCTCA 2289
                 :::   |||         |||         ||||||      :::
Db         92 nTyrGlnThrGlySer---ValValGluGlyProLysGlyGluThr-------------- 106

Qy       2288 ACTAGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCAGAGTCCCCGCATTC 2229
                                                 |||      |||
Db        107 ---------------------------------GlyProArgGlyGluArgGlyProPr 115

Qy       2228 CAGTGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGG 2169
                 |||            |||:::|||
Db        115 oGlyAlaProGlyArgAspGlyIleProGly---------------------------- 125

Qy       2168 CCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCCCCAGCCCCCAGC 2109
                 |||   |||   |||                        |||   |||   ||||||
Db        126 ----GlnProGlyIlePro--------------------GlyProProGlyProProGl 137

Qy       2108 TGTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAG 2049
                                |||      ||||||:::                |||
Db        137 y---------------ProAlaGlyLeuGlyGlyAsn---------------PheAlaPr 147

Qy       2048 CAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGG 1989
                 ::::::                     |||   :::   |||   :::||
Db        147 oGlnMetSerTyrGlyTyrAspGluLysSerAlaGlyIleSerMetProGlyProMetGl 167

Qy       1988 ACCCAGTGAGGCAGGCCCTCCACCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGG 1929
                 ||||            ||||||      :::||||||
Db        167 yProMetGlyProArgGlyProProGlyProSerGlySer------------------- 180

Qy       1928 CCAAGTCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGC 1869
                                                     |||   |||      |||
Db        181 -------------------------------------ProGly--ProGln---Gly 185
```

```
Qy     2501 -----------------------------------------GGGCTGAGGGGACCTGGT 2484
                                                      |||:::  ||||||
Db       85 PheCysProGluGluTyrValSerProAspAlaGluValIleGlyValGluGlyPro--- 103

Qy     2483 TCTTGTGTGTTGCCCCTCAGGACTCTTCCCCTACAAATAAGTCATATGTTCAAATCCCAT 2424
                                                                 |||
Db      104 ---------------------------------------------------LysGlyAsp 106

Qy     2423 GGAGGAGTGTTTCATCCTAGAAACTCCCATGCAAGAGCTACATTAAACGAAGCTGCAGGT 2364
              |||        ||||||   |||:::
Db      107 ProGlyProGlnGlyProArgGly-ProValGlyPro---------------------- 118

Qy     2363 TAAGGGGCTTAGAGATGGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCT-- 2306
               |||   :::||||||  ||||||                |||      |||
Db      119 -ProGly---GlnAspGlyIleProGlyGlnProGlyLeuProGlyProProGlyProPr 137

Qy     2305 ----TCTCTAGGTGTGTCTCAACTAGGAGGCTAGCTGTTAACC----------------- 2267
                    |||       |||||||||        :::
Db      137 oGlyProProGlyProLeuGlyLeuGlyGlyAsnPheAlaSerGlnMetSerTyrGlyTy 157

Qy     2266 ------------------CTGAGCCTGGGTAATCCACCTGCAGAGTCCCCGCATTCCAG 2226
                              :::|||:::      |||        |||
Db      157 rAspGluLysSerAlaGlyValSerValProGlyProMetGlyProSer----------- 173

Qy     2225 TGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGCCT 2166
                       |||   |||||||||
Db      174 -------GlyProArgGlyLeuPro----------------------------------- 179

Qy     2165 CCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCCCCAGCCCCCAGCTGT 2106
                       |||   |||                     |||   |||   |||
Db      180 -------GlyProPro---------------------GlyAlaProGlyPro-------- 187

Qy     2105 GCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGAGCCACATTACTTTGGCAGCAA 2046
                                     ||||||        ||||||        |||   ::
Db      188 ------------GlnGlyPheGlnGlyProProGlyGluProGlyGluProGlyGlySe 203

Qy     2045 CAGAAACTGGCGGCCAGCCCGGCAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGGA-- 1988
            :              |||        |||           :::::|||:::    |||
Db      203 rGlyProMetGlyProProGlyProProGlyProProGlyLysAsnGlyAspAspGlyGl 223

Qy     1987 ------------------CCCAGTGAGGCAGGCCCTCCA-------------------- 1967
                              |||   |||   |||||||||
Db      223 uAlaGlyLysProGlyArgProGlyGluArgGlyProProGlyProGlnGlyAlaArgGl 243

Qy     1966 ----CCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCAA 1911
                |||      ||||||                     :::   |||
Db      243 yLeuProGlyThrAlaGly---------------LeuProGlyMetLysGlyHisArgGl 258

Qy     1910 ATACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCCTGCGGCAGACACCATAT 1851
            :::                         ||||||       |||        :::
Db      258 yPheSerGlyLeuAspGlyAlaLysGlyAspThrGlyProAlaGlyProLysGlyGluPr 278

Qy     1850 AGGCAGTGA----CAGACTGGCTGAGCTGGACAATGGAGCCCATAAACAGGGATG----- 1800
            ||||||     :::   |||        ||||||        |||   |||:::
Db      278 oGlySerProGlyGluAsnGlyThrProGlyGlnMetGlyPro---ArgGlyLeuProGl 297

Qy     1799 ----------------GGGCCACCTGGGACAGCAGGAAGGCACTATCCAGGATGGCGAGG 1756
                            ||||||||||||||||||||            ||||||   |
Db      297 yGluArgGlyArgProGlyProProGlyThrAlaGlyAlaArgGly-AsnAspGlyAlaV 317

Qy     1755 TCCAGGCAGATGCCCCGGCCCGGAACCACCCTGGCCTCGGTGGGCTCACCCACCACCACA 1696
            ||   |||         ||| ||||||   |||                 |||   |||
Db      317 alGlyAlaAlaGlyPro-ProGlyProThr---GlyProThrGlyProProGlyPhePro 335

Qy     1695 CGTACGGAGACATCACAGGCAGAGGCC---CCGCAGAGCGCGGGTGGAGGT--------- 1648
                              ::::::  ||||||   ||||||   |||   |||
Db      336 GlyAlaAlaGlyAlaLysGlyGluAlaGlyProGlnGlyAlaArgGlySerGluGlyPro 355

Qy     1647 ---GGGAGCAGG------CCACTGCCTCCAGCACCCACGTGTCCATTAGGG---AAGGGA 1600
               |||   |||      |||   ||||||   |||                |||   |||
Db      356 GlnGlyValArgGlyGluProGlyProProGlyProAlaGlyAlaAlaGlyProAlaGly 375

Qy     1599 GCTCCAGGCTTAGGG------CCTGGCAGGAAGCTGGTCATCAGGCTGTCCTCACTGCTA 1546
                 ||||||         ||||||   |||       ::::::|||   |||
Db      376 AsnProGlyAlaAspGlyGlnProGlyAlaLys-GlyAlaAsnGlyAlaPro-------- 392

Qy     1545 GCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGTAGAGG 1486
                               ||||||   ||||||   |||      ||||||   |||||
Db      393 ------------------GlyIle---AlaGlyAlaProGlyPheProGlyAlaArgGl 405

Qy     1485 GAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCT 1426
            |                 :::|||   ||||||      |||   ||||||       |||
Db      405 y-----------ProSerGlyPro-GlnGlyProSerGlyAlaProGlyProLysGlyT 421

Qy     1425 GAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCACCGGCAGCCACAGGGAAA 1366
              |||         |||         |||          ||||||:::    :::|||::
Db      421 hrSerGlyGluProGly---AlaProGlyAsnLysGlyAspThrGlyAla--LysGlyGl 439

Qy     1365 GCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCTGCACCAGCCGGT----- 1311
            :                                             |||||||||
Db      439 u-------------------------------------ProGlyProAlaGlyValGl 446

Qy     1310 -------CCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACA-------- 1266
                   |||   |||              |||:::           |||
Db      446 nGlyProProGlyProAlaGlyGluGluGlyLysArgGlyAlaArgGlyGluProGlyPr 466

Qy     1265 -------GCCCCAGGCTGCCCATCCGAACGC--------------------------CT 1240
                   ||||||   |||   ||||||
Db      466 oSerGlyLeuProGlyProProGlyGluArgGlyGlyProGlySerArgGlyPheProGl 486

Qy     1239 TCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTACAGC 1180
                   |||:::||||||:::    |||:::                 ||||||   ||
Db      486 yAlaAspGlyValAlaGlyProLysGlyProSerGlyGluArgGlySerProGlyProAl 506

Qy     1179 CCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGCTGCAC 1120
            |        :::|||
Db      506 aGlyProLysGlySerPro----------------------------------------- 512

Qy     1119 AGCTCAGCCACGAAGAGCCGGCGCAGGGTGCGGGGCATGCGGCAGCACAGCTGGTGCAGC 1060
                                                              ||||||
Db      513 -----------------------------------------GlyGluAlaGlyArgPr 518

Qy     1059 CGGGGAAGCAGGGCGCCCAGGT------------TCCGGAAAGCCAAGCGGGCCCGGCAT 1012
             ||||||||||||   ||||||           :::|||   |||     ||||||
Db      518 oGlyGluAlaGlyLeuProGlyAlaLysGlyLeuThrGlySerProGlySerProGlyPr 538

Qy     1011 GGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCC 952
             |||      |||   |||                |||               |||   |||||
Db      538 oAsp------GlyLysThrGlyProProGlyProAlaGlyGlnAspGlyArgProGlyPr 556

Qy      951 AGCGCTGCCT-----------------------------------CCTCAGCCACCAGC 928
            ||||   |||                                   |||   |||
Db      556 oAlaGlyProProGlyAlaArgGlyGlnAlaGlyValMetGlyPheProGlyProLysGl 576

Qy      927 AGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGCACTCCT-- 870
                                        |||      |||   ||||||:::||||||   |||
Db      576 yThrAla--------------GlyGluProGlyLysAlaGlyGluArgGlyValProGl 591

Qy      869 ----------------------------------------------CCTGGGTGCCC 859
                                                             |||   ||
Db      591 yProProGlyAlaValGlyProAlaGlyLysAspGlyGluAlaGlyAlaGlnGlyAlaPr 611

Qy      858 AGGT---AGGGGGCCAGGGCACTGGTGTCCCAGTCAATGGCAGGCAGGA----------- 813
            ||||      ||||||                  |||         ||||||
Db      611 oGlyProAlaGlyProAlaGlyGluArgGlyGluGlnGlyProAlaGlySerProGlyPh 631

Qy      812 -------------GGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCATAGACAGAG 766
                          |||      |||   |||   :::                   |||
Db      631 eGlnGlyLeuProGlyProAlaGlyProProGlyGluAlaGlyLysProGlyGluGlnGl 651

Qy      765 TAGGCCTGGCGACAGTGGT---CCGGGTCCC------GGAAGAGGTCAGAGAGCAGGGCC 715
                  |||||||||   |||   |||||||||         ||||||:::|||   |||
Db      651 yValProGlyAspLeuGlyAlaProGlyProSerGlyAlaArgGlyGluArg---GlyPh 670

Qy      714 TCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGCCCAGGATGAGC 655
             |||   |||                |||            ||||||   ||||||
```

```
Qy     1149 GTGAAGGTCATGAGTGCCA---------------TCCAGCTGCACAGCTCAGCCACGAAG 1105
                       ||||||                  :::          ||||||
Db      296 ------------ValProProGlyGlyMetGlyGlyThrGlyGlyThrGlyProArgSe 311

Qy     1104 AGCCGGCGCAGGGTGCGGGGCA------TGCGGCAGCACAGCTGGTGCAGCCGGG----- 1056
                ||||||||||| |||      |||     |||     ||||||
Db      311 rGluGlyAlaGlyCysProAlaArgGlyAlaGlyProProAlaTrpGlyAlaGlyProPr 331

Qy     1055 -------GAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGC------------------GG 1021
                    :::   |||     |||  |||   |||                    ||
Db      331 oArgArgAspGlyGlyAsnArgGlyAspGlyAlaProGluArgGlySerGlyLeuProGl 351

Qy     1020 GCCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCC------------- 975
            |||||||    |||   |||            |||   |||   |||
Db      351 yProGlyCysGlySerProGlyLeuGlySerArgSerProProAlaGlyTrpGlyGluPr 371

Qy      974 ------------------------CTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCT-- 942
                                          ||||||  |||        :::|||
Db      371 oGlyGlyArgGlyProGlyAlaArgGluArgAlaAlaArgProGlyValArgValProAr 391

Qy      941 ---------CCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTG 892
                      |||  ||||||                                |||
Db      391 gProGlyGluProValProProGly----------------------------GlyMe 401

Qy      891 AGCAGGCCAAAGAGGCACTCCTCCT---------------GGGTGCCCAGGTAGGGGGCC 847
               |||     ||||||    |||               |||||||||  ||||||
Db      401 tGlyGlyThrGlyGlyThrGlyProArgSerGluGlyAlaGlyCysProAlaArgGlyAl 421

Qy      846 AGGGCACTGGTGTCCCAGTCAATGGCAGGCAGGA-----------------------GG 811
              |||         |||     |||   ||||||
Db      421 aGly---------ProProAlaTrpGlyAlaGlyProProArgArgAspGlyGlyAsnAr 438

Qy      810 TAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGCGACAG 751
                  |||:::|||:::    |||                        |||
Db      438 gGlyAspGlyAlaProGluArgGlySer-----------------------GlyLeuPr 450

Qy      750 TGGTCCGGGT---------CCCGGAAGAGGTCAGAGAGCAGGGCCTCCAGTGGAGTGAAG 700
              |||||||||          ||||||  |||   |||:::   ||||||
Db      450 oGlyProGlyCysGlySerProGlyLeuGlySerArgSer---ProProAlaGlyTrpGl 469

Qy      699 CACACCTGGCCACAGAAGTCCAGCA------------GCCCCACGCCCAGGATGAGCAGT 652
               ||||||  |||  |||                  |||  |||||||||     ||
Db      469 yGluProGlyGlyArgGlyProGlyAlaArgGluArgAlaAlaArgProGlyValArgVa 489

Qy      651 GCCAGCTCCAGGGGCCTGGGATCCG---------------GGCACAGCAGCCCTGCTAGC 607
            ||||  ||||||         |||                 ||||||
Db      489 lProArgProGlyGluProValProProGlyGlyMetGlyGlyThrGlyGly-------- 506

Qy      606 CAGCÇGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCAAGGACAGTGCCCAGATG 547
                          |||          ||||||||||||
Db      507 ----------ThrGlyProArgSerGluGlyAlaGlyCysProAlaArgGlyAlaGlyPr 523

Qy      546 AAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTAGGAGCGGG 487
               |||   |||                     |||               ||||||
Db      523 oProAlaTrpGlyAlaGlyProProArgArgAspGlyGlyAsnArgGlyAspGlyAlaPr 543

Qy      486 ACACAGA-----------CCAGGCCCAGCACTGGACCAATGCCCAGCACCATGGTCATG 439
               |||           |||||||||           |||    :::|||
Db      543 oGluArgGlySerGlyLeuProGlyProGly---------CysGlySerPro---GlyLe 559

Qy      438 AACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCC 379
               |||     ||||||          |||       |||
Db      559 uGlySerArgSerProProAlaGlyTrpGlyGluProGlyGlyArg-------------- 574

Qy      378 AAACACACCTCCAGGCCAAAGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGTGC 319
                      ||||||  |||            ||||||  |||:::       ||||
Db      575 -------GlyProGlyAlaArgGluArgAla------AlaArgProGlyValArg-ValP 590

Qy      318 CGCAGCAGGCGGCTCACCCACAGCCTCTGGACCATAGTGGGCCAGGCGGGTAGGGCTCAG 259
            ||   :::        |||   |||               ||||||      |||    |
Db      590 ro---ArgProGlyGluProValPro---------------ProGly------GlyMetG 602

Qy      258 GGGGCCGTTCAGGCACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCT 199
             |||||    ||||||                                  |||     |||
Db      602 lyGlyThrGlyGlyThrGlyProArgSerGluGlyAlaGlyCysProAlaArgGlyAlaG 622

Qy      198 CTCCTCCTTGCTGCCGCCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCT 139
               ||||||           |||      :::   |||    :::          :::
Db      622 lyProProAlaTrpGlyAlaGlyProProArgArgAspGlyGlyAsn------ArgGlyA 640

Qy      138 TTGGTGCC-----------------------GGTCCAGCTTCTCAGCCCATGCTCAACA 103
               ||||||                         ||||||
Db      640 spGlyAlaProGluArgGlySerGlyLeuProGlyProGly------------------ 653

Qy      102 CCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTCA 65
                     ||||||:::|||       |||       |||
Db      654 --------CysGlySerProGlyLeuGlySerGlnSer 663
```

RESULT 13
Q63079

```
ID   Q63079      PRELIMINARY;      PRT;  1453 AA.
AC   Q63079;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Collagen alpha1 (Fragment).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE OF 1-1092 FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=TOOTH;
RA   Brandsten C., Lundmark C., Christersson C., Hammarstrom L., Wurtz T.;
RT   "Expression of Collagen alpha1(I) mRNA variants during Tooth and Bone
RT   Formation in the Rat.";
RL   J. Dent. Res. 0:0-0(0).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=TOOTH;
RA   Wurtz T.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z78279; CAB01633.1; -.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR000885; Fib_collagen_C.
DR   InterPro; IPR001007; VWF_C.
DR   Pfam; PF01410; COLFI; 1.
DR   Pfam; PF01391; Collagen; 18.
DR   ProDom; PD000007; Collagen; 3.
DR   ProDom; PD002078; Fib_collagen_C; 1.
DR   SMART; SM00038; COLFI; 1.
DR   SMART; SM00214; VWC; 1.
DR   PROSITE; PS01208; VWFC; 1.
KW   Collagen.
FT   NON_TER      1      1
SQ   SEQUENCE   1453 AA;  137887 MW;  E6896BDC19A4A1D8 CRC64;

Alignment Scores:
Pred. No.:              7.32e-15      Length:         1453
Score:                  357.00        Matches:        283
Percent Similarity:     31.84%        Conservative:   50
Best Local Similarity:  27.06%        Mismatches:     362
Query Match:            5.73%         Indels:         352
DB:                     11            Gaps:           59

US-09-759-143-110 (1-3410) x Q63079 (1-1453)

Qy     2594 TGCCTGTGT---------------CTCTGTGATGGC---------AACAGAAGGACCAAC 2559
              |||:::|||                :::|||||||||                  :::
Db       52 CysIleCysHisAsnGlyThrAlaValCysAspGlyValLeuCysLysGluAspLeuAsp 71

Qy     2558 AGGCCACATCCTGATAAAAGGTAAGAGGGGGGTGGATCAGCAAAAAGACAGTGCTGT--- 2502
                |||:::|||   ||||||   ||||||                     :::||||||
Db       72 CysProAsnProGlnLysArg---GluGly-----------------GluCysCysPro 84
```

```
Db        446 ThrGlyGlyHisProAlaAlaPro---------GlyAlaProGlyProArgSerProArg 462

Qy        410 GCAGAGGCGGCACATAGGTGATGCCTGCGGCCAAACACACCTCCA------GGCCAAAGG 357
                 |||                     |||   :::  ||||||      |||||||||
Db        463 ThrGluArgArgArgGlySerAlaGlnArgGlyHisProProProGlyAlaGlyGlnArg 482

Qy        356 -----------TTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGTGCCGCAG----- 314
                         |||  ||||||      |||      ||| |||:::
Db        483 ProSerGlyProThrGlyGlyHisProAlaAlaProGlyAlaProGly-ProArgSerPr 502

Qy        313 --------------------------CAGGCGGCTCACCCACAGCCTCTGGACCATAG 282
                                        |||      ||||||  |||
Db        502 oArgThrGluArgArgArgGlySerAlaGlnArgGlyHisProProProGlyAlaGlyGl 522

Qy        281 TGGGCCAGGCGGGTAGGGCTCAGGGGGCCGTTCAGGCACTCCAGAACTG---------- 233
                 |||  |||      :::||||||            |||
Db        522 nArgProSerGly---ProThrGlyGlyHisProAlaAlaProGlyAlaProGlyProAr 541

Qy        232 ---------------CTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTG 189
                             |||  |||||||||   :::          ||||||
Db        541 gSerProArgThrGluArgArgArgGlySerAla---GlnArgGlyHisProProProGl 560

Qy        188 CTGCCGCCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCC---------- 140
                    |||  |||              |||  |||        |||
Db        560 yAlaGlyGlnArgProSerGlyProThrGlyGlyHisProAlaAlaProGlyAlaProGl 580

Qy        139 ---------------TTTGGTGCCGGTCCAGCTTCTCAGCCC 113
                             |||:::|||||||||      |||
Db        580 yProProAsnProGluArgGlySerGlyProAlaAspProPro 594
```

RESULT 12
O41973

```
ID   O41973      PRELIMINARY;      PRT;    727 AA.
AC   O41973;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   Hypothetical 69.4 kDa protein.
GN   GAMMAHV.M10C.
OS   murid herpesvirus 4.
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae.
OX   NCBI_TaxID=33708;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WUMS;
RX   MEDLINE=97366649; PubMed=9223479;
RA   Virgin H.W. IV, Latreille P., Wamsley P., Hallsworth K., Weck K.E.,
RA   Dal Canto A.J., Speck S.H.;
RT   "Complete sequence and genomic analysis of murine gammaherpesvirus
RT   68.";
RL   J. Virol. 71:5894-5904(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WUMS;
RA   Latreille P., Wamsley P., Waterston R.H.;
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U97553; AAB66422.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   727 AA;  69448 MW;  5FFDB077864B08D1 CRC64;

Alignment Scores:
Pred. No.:              4.86e-15     Length:        727
Score:                  359.00       Matches:       229
Percent Similarity:     31.50%       Conservative:  28
Best Local Similarity:  28.06%       Mismatches:    270
Query Match:            5.77%        Indels:        289
DB:                     12           Gaps:          49

US-09-759-143-110 (1-3410) x O41973 (1-727)

Qy       2181 CCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGC 2122
                ||||||   :::                                 |||||||||   |||
Db         30 ProProArgArgAspGlyGlyAsnArgGlyAspGlyAlaProGluArgGlySerGly-Le 49

Qy       2121 CCCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGAGCCAC 2062
               |||   |||   |||   :::         |||   |||
Db         49 uProGlyProGlyCysGlySerProGlyLeuGlySer---------------------- 61

Qy       2061 ATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGG----------- 2013
                                         ||||||  |||   |||
Db         62 ---------------------------ArgSerProProAlaGlyTrpGlyGluProGl 72

Qy       2012 ------------GGCTAACAGGAGCGGGGAGCT------GGGACCCAGTGAGGCAGGCC 1972
                          |||   :::||||||   |||    |||    :::       ||||
Db         72 yGlyArgGlyProGlyAlaArgGluArgAlaAlaArgProGlyValArgValProArgPr 92

Qy       1971 CTCCACCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCTTGTCA 1912
               |    |||:::
Db         92 oGlyGluProVal----------------------------------------------- 96

Qy       1911 AATACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCCTGCGGCAGACA----- 1857
                                           |||    |||:::        |||
Db         97 -----------------------------ProProGlyGlyMetGlyGlyThrGlyGl 106

Qy       1856 -------------------------------CCATATAGGCAGTGACAGACTGGCTGAGC 1828
                                             |||   |||
Db        106 yThrGlyProArgSerGluGlyAlaGlyCysProAlaArg------------------- 119

Qy       1827 TGGACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCACTATCC 1768
                |||     .     |||      |||   |||||||||          ||| |||
Db        120 -GlyAlaGlyProProAlaTrpGlyAlaGlyProProArgArgAspGly-GlyAsnArgG 139

Qy       1767 AGGATGGCGAGGTCCAGGCAGATGCC------------CCGGCCCGGAACCACCCTGGCC 1720
               ||||||   .   :::    .   :::          |||         ||||||
Db        139 lyAspGlyAlaProGluArgGlySerGlyLeuProGlyProGlyCysGlySerProGlyL 159

Qy       1719 TCGGTGGGCTCACCCACCACCACACGTACGGAGACATCACAGGCAGAGGCCCCGCAGAGC 1660
               |||||      :::          :::|||:::       |||||||||||       |
Db        159 euGlySerArgSerProProAlaGlyTrpGlyGluProGlyGlyArgGlyProGlyAlaA 179

Qy       1659 GC-----------------GGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACG 1618
               ||                |||   |||      ||||||    |||   |||
Db        179 rgGluArgAlaAlaArgProGlyValArg-ValProArgProGlyGluProValProPro 198

Qy       1617 TGTCCATTAGGGAAGGGAGCTCCAGGCTTAGGGCCTGGCAGGAAGCTGGTCATCAGGCTG 1558
                  :::|||    |||        |||   ||||||  ||| |||      :::|||||
Db        199 GlyGlyMetGly---GlyThrGlyGlyThrGlyPro---Arg-SerGluGlyAlaGlyCy 216

Qy       1557 TCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTCC 1498
               ||||               |||           ||||||  |||
Db        216 sProAla------ArgGlyAlaGlyProProAlaTrpGlyAlaGlyProProArgArgAs 234

Qy       1497 CGGTGGT---AGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGAAC 1441
                ||||||   ||||||                   ||||||   ||||||
Db        234 pGlyGlyAsnArgGlyAspGlyAlaProGluArgGlySer---GlyLeu----------- 249

Qy       1440 CCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACA---GGCATGTGGCA 1384
                                       |||   |||   ||||||:::   |||:::
Db        250 -------------------------ProGlyProGlyCysGlySerProGlyLeuGlySe 261

Qy       1383 CCGGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGC 1324
               |||   |||
Db        261 rArgSerPro------------------------------------------------- 264

Qy       1323 TGCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGG---CGCACTGCAGGAAC 1267
                    |||||||||                         |||         |||
Db        265 ----ProAlaGly--------------------------TrpGlyGluProGlyGlyAr 274

Qy       1266 AGCCCCAGGCTGCCCATCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGC 1207
                 ||||||      ||||||
Db        274 gGlyProGlyAla---ArgGluArgAla----------------------------- 282

Qy       1206 TCAGCTCTGGGCACGCCCTGGTACA---GCCCCTCGCCCACGAAATCCGTGTAAAACAGC 1150
                        ||||||||||||          |||||||||      |||
Db        283 ---------AlaArgProGlyValArgValProArgProGlyGluPro----------- 295
```

```
Qy        166 CAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCAGCTTCTCAGCCCATGCTC 107
                  |||||||||                    |||||||||     |||
Db        961 SerProGlyAla--------------------ValGlyProAlaGlyAsnPro------ 971

Qy        106 AACACCTGCTGCTGTGGGGCACCTCAGTGGGGACACGTCTCATCACTCAGATCCTGGC-- 49
                 ||||||           :::   |||           :::     ||||||
Db        972 Gly-ProAlaGlyGluAsnGlyMetProGlySerAspGlyAsnAspGlyAlaProGlyPr 991

Qy         48 ----------CGAGGCGCGCGGCTGTCACCCGGAGCCAGCGCGTGCAGGCTGGTTCCC 1
                        ||||||    :::          |||||||||          |||
Db        991 oGlnGlySerArgGlyGluLysGlyAspThrGlyAlaSerGlyAlaAsnGlySerPro 1010

RESULT 11
Q9Q5K9
ID   Q9Q5K9      PRELIMINARY;      PRT;   608 AA.
AC   Q9Q5K9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   NTR.
OS   Herpesvirus papio.
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae; Lymphocryptovirus.
OX   NCBI_TaxID=10394;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BABOON LYMPHOCRYPTOVIRUS BA65;
RA   Zong J.-C., Ryan J., Ling P.D., Loeb D.D., Pagano J.S., Hayward S.D.,
RA   Hayward G.S.;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF200364; AAF23950.1; -.
SQ   SEQUENCE   608 AA;  60458 MW;  4BF82ACB0A029627 CRC64;

Alignment Scores:
Pred. No.:              3.51e-15       Length:          608
Score:                  361.00         Matches:         203
Percent Similarity:     35.05%         Conservative:    28
Best Local Similarity:  30.80%         Mismatches:      263
Query Match:            5.80%          Indels:          165
DB:                     12             Gaps:            36

US-09-759-143-110 (1-3410) x Q9Q5K9 (1-608)

Qy       1805 GGGATGGGGCCACCTGGGACA---------------GCAGGAAGGCACTATCCAGGATGG 1761
                 |||   |||   |||||||||               |||||||||
Db         11 GlyProGlyArgProGlyThrPheProGlyArgGluAlaGlyArgGluAlaGlyArgThr 30

Qy       1760 CGAGGTCCAGGCAGATGCCCCGGCCCGGAACCACCCTGGCCTCGGTGGGCTCACCCACCA 1701
                 ||||||          ||||||   |||   ||||||           |||   |||
Db         31 TrpGlyPro------CysProAlaProSerProProProAlaPheGlnValHisGlyPro 48

Qy       1700 CCACACGTACGGAGACATCACAGGCAGAGGCCCCGCAGAGCGCGGGT-----------G 1653
                          |||         ||| ||||||   |||:::||| |||           |
Db         49 ArgSerProArgThrGluArgArg-ArgGlySerAlaGlnArg-GlyHisProProProG 68

Qy       1652 GAGGTGGGAGCAGGCCACTGCCTCCAGCACCCACGTGTCCATTAGGGAAGGGAGCTCCAG 1593
                 ||   |||   ||||||      |||            |||         |||||||||
Db         68 lyAlaGlyGlnArgProSerGlyProThrGlyGlyHisProAlaAlaProGlyAlaProG 88

Qy       1592 GC---TTAGGGCCTGGCA---------------GGAAGCTGGTCATCAGGCTGTCCTCAC 1551
              ||      |||                          ||||||      |||   |||
Db         88 lyProArgSerPro-ArgThrGluArgArgArgGlySerAlaGlnArgGlyHisProPro 107

Qy       1550 TGCTAGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCTTCTCCCGGTGGT 1491
                             |||   |||   |||         |||   ||||||:::||||||
Db        108 ProGlyAlaGlyGlnArgProSerGlyProThrGlyGlyHisProAlaAlaProGlyAla 127

Qy       1490 AGAGGG------------AGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGGTGA 1443
                 |||               |||         |||   |||||||||
Db        128 ProGlyProArgSerProArgThrGluArgArgArgGlySerAla--------------- 142

Qy       1442 ACCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACAGGCATGTGGCAC 1383
                                       :::          |||      |||      |||
Db        143 ---------------GlnArgGlyHisProProProGlyAlaGly-------------- 152

Qy       1382 CGGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGAATCGCT 1323
              ::::::|||   |||                   |||                   |||
Db        153 GlnArgProSerGlyProThrGlyGlyHisPro-----------------------Ala 164

Qy       1322 GCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGGCGCACTGCAGGAACAGCC 1263
              ||||||      |||   |||   |||   |||            |||   |||:::|||
Db        165 AlaProGlyAlaProGlyPro---ArgSerProArgThrGluArgArgArgGlySerAla 183

Qy       1262 CCAGGCTGCCCATCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGCCCGGCTCAG 1203
               |||   ||||||                    ||||||   |||   |||   ||
Db        184 -GlnArgGlyHisPro-----------------ProProGlyAlaGlyGlnArgProSe 197

Qy       1202 C---------TCTGGGCAC------GCCCTGGTACAGCCCCTCGCCCACGAAATCCGTGT 1158
              |               ||||||      |||         |||            |||
Db        197 rGlyProThrGlyGlyHisProAlaAlaProGlyAlaProGlyProArgSerProArgTh 217

Qy       1157 AAAACAGCGTGAAGGTCATGAGTGCCATCCAGCTGCACAGCTCAGCCACGAAGAGCCGGC 1098
               ::::::|||   |||                  :::||||||      |||   |||
Db        217 rGluArgArgArgGly---------------SerAlaGlnArgGlyHisProProProGl 232

Qy       1097 GCAGGGT--------------------GCGGGGCATGCGGCAGCACAGCTGGTGCAGCC 1059
                 |||                        ||||||   ||||||             ||
Db        232 yAlaGlyGlnArgProSerGlyProThrGlyGlyHisProAlaAla------------Pr 248

Qy       1058 GGGGAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGCGGGCCCGGCATGGACAGCAGTGGG 999
              ||||      ||| ||||||      :::   :::|||   |||           |||   |
Db        248 oGlyAlaProGly-ProArgSerProArgThrGluArgArgArgGlySerAlaGlnArgG 268

Qy        998 GC---GACAAGGAGGGGGCCGACAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCT 942
              ||         ||||||            :::|||      |||         |||:
Db        268 lyHisProProProGlyAlaGlyGlnArgProSerGlyProThrGlyGlyHisProAlaA 288

Qy        941 CCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAA 882
              ::                      :::   :::      :::   |||   |||   |||
Db        288 laProGlyAlaProGlyProArgSerProArgThrGluArgArgArgGlySer-AlaGln 307

Qy        881 AGAGGCACTCCTCCTGGGTGCCCAGGTAGGGGGCCA--------------GGGCACTGG 837
              ||||||   ||||||         |||:::   |||                  ||||||
Db        308 ArgGlyHisProProProGlyAlaGlyGlnArgProSerGlyProThrGlyGlyHis--- 326

Qy        836 TGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGG 777
                 |||:::          |||   |||:::||||||   |||:::
Db        327 ---ProAla---------AlaProGlyAlaProGly---ProArgSerProArgThrGlu 341

Qy        776 CATAGACAGAGTAGGCCTGGCGACAGTGGTCCGGGTCCCGGAAGAGGTCAGAGA------ 723
                 |||:::                            |||   ||||||   |||||||||
Db        342 ArgArgArgGlySerAlaGlnArgGlyHisProProProGlyAlaGlyGlnArgProSer 361

Qy        722 ---------GCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCC 672
                             |||   |||                ||||||   |||||||||
Db        362 GlyProThrGlyGlyHisProAlaAlaProGlyAlaProGlyProArgSerProArgThr 381

Qy        671 CCACGCCCAGGATGAGCAGTG-----------CCAGCTCCAGGGGCCTGGGATCCGGGC 624
                 |||            :::            |||   |||||||||
Db        382 GluArgArgArgGlySerAlaGlnArgGlyHisProProProGlyAla----------- 397

Qy        623 ACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGATGC--- 567
                                     ||||||   |||         |||
Db        398 ---------------GlyGlnArgProSerGlyProThrGlyGlyHisProAlaAlaPro 412

Qy        566 ---------------------CCAAGGACAGTGCCCAGATGAAGGGCCGGCGGCGGCCAT 528
                                   |||||||||                |||   |||
Db        413 GlyAlaProGlyProArgSerProArgThr------GluArgArgArgGlySerAlaGln 430

Qy        527 AGCGTCCACGCCAGTGGTCACTGGCTGAGCCTAGGAGCGGGACACAGACCA---GGCCCA 471
                 |||                              |||||||||   ||||||   ||||||
Db        431 ArgGlyHisProProPro---------------GlyAlaGlyGlnArgProSerGlyPro 445

Qy        470 GCACTGGACCAATGCCCAGCACCATGGTCATGAACTTCTCCTCTACCCCCACTTCCAGCA 411
                             ||||||                :::|||   |||      |||
```

```
                 ||||||              ||||||  |||||||||     |||
Db     399 lyAsnProGlyGlnProGlyAlaGlnGlyGluSerGlyProLeuGlyProArgGlyGluT 419

Qy    1977 CAGGCCCT------CCACCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGCCAAG 1924
             ||||||    ||||||
Db     419 hrGlyProAlaGlyProPro---------------------------------------- 425

Qy    1923 TCGCTCTTGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCCTGCG 1864
                                               |||:::      :::
Db     426 ---------------------------GlyAlaGlnGlyGluSerGlyGluArgGlyS 436

Qy    1863 GCAGACACCATATAGGCAGTGA----CAGACTGGCTGAGCTGGACAATGGAGCCCATAAA 1808
             |||      :::|||            |||      |||:::      |||
Db     436 erArgGlyAlaLeuGlyProAlaGlyProProGlyGlyValGlyGluArgGlyPro---- 454

Qy    1807 CAGGGATGGGGCCACCTGGGACAGCAGGA----------------------------A 1778
                ||||||||||||||   :::|||                             |
Db     455 -----MetGlyProProGlyMetSerGlyAlaProGlyAlaProGlyAlaLysGlyAspA 473

Qy    1777 GGCACTATCCAGGATGGCGAGGT----------------------------CCAGGCA 1748
           ||      ||||||  ||||||                              ||||
Db     473 rgGlyLeuProGlyGluArgGlySerAlaGlySerLysGlySerAlaGlyGluSerGlyA 493

Qy    1747 GATGCCCCGGCCCGGAACCACCCTGGCCTCGGTGGGCTCACCCACCACCACACGTACGGA 1688
           ||          |||      ||||||     |||||||||
Db     493 rgProGlyGluPro-GlyMetProGlyGlnArgGlyLeuThr------------------ 506

Qy    1687 GACATCACAGGCAGAGGCCCCGCAGAGCGCGGGTGGAGGTGGGAGCAGGCCACTGCCT-- 1630
                             ||||||     :::     |||        :::|||    |||
Db     507 --------------GlyProProGlyLys-GlnGlyArgAspGlyLysProGlyProAl 521

Qy    1629 ---------------CCAGCACCCACGTGTCCATTAGGG------------AAGGGAGC 1598
                          |||     :::   |||   |||              :::|||
Db     521 aGlyAlaProGlyGluProGlyAsnSerGlyProAlaGlyAlaSerGlyGlnArgGlyLe 541

Qy    1597 TCCAGGCTTA------------GGGCCT----------------------------GG 1580
             |||||||||             ||||||                             ||
Db     541 uProGlyLeuValGlyLeuProGlyProGlnGlyGlnArgGlyGluArgGlyGluAspGl 561

Qy    1579 CAGGAAGCTGGTCATCAGGCTGTCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGTATTT 1520
           | |||     |||   |||   |||
Db     561 y-GlyGlnGlySerProGlyAlaPro-------------------------------- 569

Qy    1519 GGGCAGGAACACCTGCTTCTCCCGGTGGTAGAGGGAGGCCAGTGTGTAGGGCAGGATCTG 1460
                |||           ||||||  ||||||   |||              |||   |
Db     570 -----GlyLeuThrGlyGluProGlyLysArgGlyGluPro-----------GlyValA 584

Qy    1459 CAGGGCTGAGAAGGTGAACCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTG 1400
           |||||                                          |||   |||
Db     584 laGly-----------------------------------ProProGlyProGlnGlyS 592

Qy    1399 GGACAGGCATGTGGCACCGGCAG--CCACAGGGAAAGCTGCCACACTGGCCAAATAGAC 1343
               |||                |||   |||||||||
Db     592 erAlaGlyGluArgGlyAsnGlnGlyProGlnGly----------------------- 603

Qy    1342 TGCTCGAGTGCCGAATCGCTGCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGAT 1283
                :::                  |||   ||||||               ||||||
Db     604 -----GlnAlaGlySerMetGlyProProGlyPro---------------ProGlyAlaS 617

Qy    1282 GGCGCACTGCAGGAACAGCC------CCAGGCTGCCCATCCGAACGCCTTCATCATAGTG 1229
                   ||||||                |||    |||   |||
Db     617 erGlyAspAlaGlyAlaGlnGlyAspAsnGlyProProGlyGlu--------------- 631

Qy    1228 TCTCCGGGCCTCGGTGCCCGGCTCAGCTCTGGGCACGCCCTGGTACA------------ 1182
             |||||||||      |||                     ||||||
Db     632 --SerGlyProGluGlyPro-----------------ProGlyAlaArgGlyGluArgG 645

Qy    1181 --GCCCCTCGCCCACGAAATCCGTGTAAAACAGCGTGAAGGTCATGAGTGCCATCCAGCT 1124
             ||||||     |||   |||
Db     645 lyAlaProGlyGluArgGlyPro-------------------------GlnGlyLeuT 656

Qy    1123 GCACAGCTCAGCCACGAAGAGCC---GGCGCAGGTGCGGGGCATGCGGCAGCACAGCT- 1068
              ||||||  ||||||       |||         |||        |||:::|||

Db     656 hrGlyAlaGlnGlyArgArgGlyIleGlyGlyArgAlaGlyAsnTyrGlyAlaThrGlyP 676

Qy    1067 -----GGTGCAGCCGGGGAAGCAGGGCGCCCAGGTTCC-------------------G 1034
                 |||      ||||||   |||   ||||||:::                    |
Db     676 roProGlyGlnLysGlyGluMetGlyProProGlyAsnValGlyLeuGlnGlyProProG 696

Qy    1033 GAAAG------CCAAGCGGGCCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGA 980
           ||:::       ||||||||||||                 ||||||     |||||||||
Db     696 lyGluLeuGlyProSerGlyPro-----------ProGlyAla------ArgGlyProP 710

Qy     979 CAGCCCTTCTGCTGGCTCGGTGGGGCCCAGCGCTGCCTCCTCAGCCACCAGCAGTGTGGC 920
                                   ||||||:::     |||   |||
Db     710 ro-------------------GlyProSerGlySerProGlyProAsp---------- 719

Qy     919 TGCTACGCAGGTGAGGAAGATGAGGGTGAGCAGGCCAAAGAGGCACT------------ 873
                             |||     |||      |||   ||||||
Db     720 -------------GlyProAlaGlyAlaGluGlyAspArgGlyProValGlyProMetG 735

Qy     872 ----------CCTCCTGGGTGCCCAGGT---AGGGGG-----------CCAGGGCACT 839
                     |||   |||  ||||||  ||||||              ||||||
Db     735 lyProSerGlyProSerGlyMetProGlyGluArgGlyAspAsnGlyGluProGly---- 753

Qy     838 GGTGTCCCAGTCAATGGCAGGCAGGAGGTAGCCCAGGCAGCCCCCAAGACTGATCATGAA 779
                                ||||||:::||||||          |||
Db     754 ------------------ProGlyGlyAlaProGlyGlnArgGlyAsp---------- 763

Qy     778 GGCATAGACAGAGTAGGCCTGGCGACAGTGGTCCG------GGTCCCGGAAGAGGTCAGA 725
                               |||||||||   ||||||           ||||||
Db     764 ----------------ProGlyAspLeuGlyProGlnGlySerProGlySerProGlyP 778

Qy     724 GAGCAGGGCCTCCAGTGGAGTGAAGCACACCTGGCCACAGAAGTCCAGCAGCCCCACGC- 666
              |||||||||||          ||||||   :::                |||
Db     778 heAlaGlyProProGlyArgSerGlyAsnProGlyProGlnGlyGluLeuGlyProThrG 798

Qy     665 ----------------------CCAGGATGAGCAGTGCCAGCTCCAGGGGCCTGGGATC 629
                                 ||||||     |||:::              ||||
Db     798 lyAlaArgGlyGluThrGlyGlyProGly---------ProSerGlyProThrGlyAspP 815

Qy     628 CGGGCACAGCAGCCCTGCTAGCCAGCCGGCCCTTGGGATGAGAAAGAGGCTCAGCAGGAT 569
            |||||          |||    :::              ||||||||     |||
Db     815 roGlyProGlnGlyProLeuGlyAlaProGlyGlnGlnGlyGluArgGlyGluThrGly- 834

Qy     568 GCCCAAGGACAGTGCCCAGATGAAGGGCCGGCGGCGGCCATAGCGTCCACGCCAGTGGTC 509
                            |||:::
Db     835 -------------ProGln---------------------------------------- 836

Qy     508 ACTGGCTGAGCCTAGGAGCGGGACACAGACCAGGCCCAGCACTGGACCAATGCCCAGCAC 449
                         |||   |||      |||||||||   :::            ||||
Db     837 -------------GlyGlnGlyGlyProProGlyPro---IleGlySerLeuGlyAlaP 851

Qy     448 CATGGTCATGAACTTCTCCTCTACCCCCACTTCCAGCAGCAGAGGCGGCACATAGGTGAT 389
            ||    :::            |||   |||    |||:::     |||
Db     851 roGlyAlaGlnGlyProProGlyProThrGlyProSerGlyAsnAlaGly---------- 867

Qy     388 GCCTGCGGCCAAACACACCT--------------CCAGGCCAAAGGTTAGCAGGTTGAC 344
                  |||      |||               |||||||||       :::|||    |
Db     868 -----SerProGlyGlnProGlyAlaArgGlyGluProGlyGln------SerGlySerP 884

Qy     343 CAGCAAGAGCTGGGCTTTCCGGTGCCGCAGCA--------------------GGCGGCT 305
            ||     :::    ||||||:::|||                             |||:
Db     884 roGlyGlnProGlyLeuAlaGlyArgThrGlyProSerGlyGluArgGlyAspLysGlyA 904

Qy     304 CACCCACAGCCTCTGGACCATAGTGGG---CCAGGCGGGTAGGGCTCAGGGGGCCGTTCA 248
            ::             |||||||||   |||    ||||||     |||      |||:::|||
Db     904 snAspGlyGlnSerGlyPro-ProGlyProProGlyProAlaGlyProAlaGlyGlnSer 923

Qy     247 GGCACTCCAGAACTGCTTCGTCTC------GGCTCTGCTCCAGAAGCTGCGGCCTCTCCT 194
            |||           :::|||   |||       ||||||   |||              |||
Db     924 Gly---------IleLeuGlyLeuAlaGlyGlySerGlyProArgGlyProGlyGlyPro 940

Qy     193 ------------------------------CCTTGCTGCCGCCAACTGCCTAGGAAT 167
                                         |||        :::          |||
Db     941 AlaGlyProProGlyAlaAlaGlySerArgGlyProAlaGlyLysSerGlyAspArgGly 960
```

```
                                  |||        ||||||    ||||||
Db        644 --------------------GluArgGlySerGlyLeuProGlyPro----------- 652

Qy        239 AGAACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTCTCCTCCTTGCTGCCGCCA 180
                                                ||||||  ||||||
Db        653 -----------------------------GlyCysGly-SerPro-----------Gl 658

Qy        179 ACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTTTGGTGCCGGTCCA----- 125
                 |||         |||   |||    :::    :::|||   |||   ||||||
Db        658 yLeuGlySerArgSerProProAlaGlyTrpGlyGluProGlyGlyArgGlyProGlyAl 678

Qy        124 ---------GCTTCTCAGCCCATGCTCAACACCTGCTGCTGTGGGGCACCTCAGTGGGG 75
                       |||::::::|||    :::                 |||   |||:::  ||
Db        678 aArgGluArgAlaAlaArgProGlyValArgValProArgProGlyGluProLysArgGl 698

Qy         74 A 74
              |
Db        698 y 698
```

RESULT 10
Q26634

```
ID   Q26634      PRELIMINARY;      PRT;  1414 AA.
AC   Q26634;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Alpha-1 collagen.
GN   COLP1ALPHA.
OS   Strongylocentrotus purpuratus (Purple sea urchin).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC   Strongylocentrotus.
OX   NCBI_TaxID=7668;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92348411; PubMed=1639795;
RA   Exposito J.-Y., D'Alessio M., Solursh M., Ramirez F.;
RT   "Sea urchin collagen evolutionarily homologous to vertebrate Pro-
RT   alpha-2(I) collagen.";
RL   J. Biol. Chem. 267:15559-15562(1992).
DR   EMBL; M92040; AAA30035.1; -.
DR   InterPro; IPR000087; Collagen.
DR   InterPro; IPR000885; Fib_collagen_C.
DR   Pfam; PF01410; COLFI; 1.
DR   Pfam; PF01391; Collagen; 18.
DR   ProDom; PD000007; Collagen; 2.
DR   ProDom; PD002078; Fib_collagen_C; 1.
DR   SMART; SM00038; COLFI; 1.
KW   Collagen.
SQ   SEQUENCE   1414 AA;  133025 MW;  94D9CDA71A9FD73D CRC64;

Alignment Scores:
Pred. No.:                2.58e-15      Length:         1414
Score:                    364.00        Matches:        321
Percent Similarity:       30.88%        Conservative:   57
Best Local Similarity:    26.23%        Mismatches:     381
Query Match:              5.85%         Indels:         467
DB:                       5             Gaps:           78

US-09-759-143-110 (1-3410) x Q26634 (1-1414)

Qy       3029 GGGAAAGTTGGGGGTAGGGGAAAGTTGGGGGTAGGGGAAATTTTGGGCAGTGCCTTCATC 2970
                |||      |||    :::|||:::    |||          :::|||       ||||||
Db         43 GlyAspLysGlyGlnLysGlyGluProGlyAspAlaAspIle---AsnSerAlaAsnPhe 61

Qy       2969 AGCCCAGTCCTAGAGAGAGTAGAGGGG--------------------AGTGGAAGTGGG 2931
                   |||   |||                 |||                     ||||||
Db         62 ProProGlyLeuProGlyProValGlyProProGlyProSerGlyProSerGlyProAla 81

Qy       2930 GGGAACCAGGCT--------------GGGCCAAGA-----------GAAGAGGGGTGG 2898
                 ||||||                  |||||||||                :::|||
Db         82 GlyAsnAsnGlyProProGlyProAsnGlyProArgGlyAsnProGlyMetAspGlyLeu 101

Qy       2897 TTAGGGAAGCCGTTGAGACCTGAAGCCCCACCCTCTACCT------------------- 2858
                 |||   |||      |||      ||||||    |||
Db        102 ThrGlyLeuProGlyIlePro---GlyProProGly-ProProGlyLysSerGlySerLe 120

Qy       2857 ---------------------TCCTTCAACACCCTAACCTTGGGTAACAGCATTTGGAA 2820
                                   |||||||||              |||  |||:::
Db        120 uValAlaSerAlaGlnThrSerSerPheAsnLys---------GlyProSerLeuAlaGl 137

Qy       2819 TTATCATTTGGGATGAGTAGAATTTCCAAGGTCCTGGGTTAGGCATTTTGGGGGGCCAGA 2760
                |||              ::::::|||::::::          :::      |||   |||
Db        137 yTyr--------------GlnTyrProGlnAla------GlnAlaAlaGlyThrProGl 150

Qy       2759 CCCCAGGAGAAGAAGATTCTGGCAATGATCAGCCCAATGACCAGCTATCTCAGGGGACCT 2700
               ||||||
Db        150 yProArg----------------------------------------------------- 152

Qy       2699 GATTGTTGGGGATCCCCCACCCTACCCAAATATTAGACACCAACACAGAAAAGCTAGCAA 2640
                               ||||||       |||       |||
Db        153 ---------GlyProProGlyProProGlySerArgGlyProGlnGlyLeuThrGlyPr 169

Qy       2639 TGGATTCCCTTCTACTTTGTTAAATAAATAAGTTAAATATTTAAATGCCTGTGTCTCTGT 2580
                    ||||||
Db        169 oSerGlyProSer------------------------------GlyGluThrGlyPr 178

Qy       2579 GATGGCAACAGAAGGACCAACAGGCCACATCCTGATAAAAGGTAAGAGGGGGGTGGATCA 2520
                          ||||||  |||          :::  |||:::   |||   |||
Db        178 oSerGlyAsnSerGlyProProGlyProSerGlyLeuProGlyArgProGlySerAspGl 198

Qy       2519 GCAAAAAGACAGTGCTGT-------GGGCTGAGGGGACCTGGTTCTTGTGTGTTGCCCCT 2467
                       :::                  |||   |||||||||
Db        198 yAspAspGlyThrProGlySerGlnGlyGlnArgGlyProAlaGlyThrProGlySerAr 218

Qy       2466 CAGGACTCTTCCCCTACAAATAAGTCATATGTTCAAATCCCATGGAGGAGTGTTTCATCC 2407
                 |||      :::      :::          |||   |||   |||:::        ||
Db        218 gGlyThrProGlyMetProGlyAlaProGlyMetLysGlyHisGlnGlyLeu------Pr 236

Qy       2406 TAGAAACTCCCATGCAAGAGCTACATTAAACGAAGCTGCAGGTTAAGGGGCTTAGAGATG 2347
              |     :::  ::::::               |||            :::|||      ||||
Db        236 oGlyMetThrGlySerLysGlyGluArgGlyGluGlyGly-GluArgGly---SerAspG 255

Qy       2346 GGAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCTCAAC 2287
               ||  ||||||               |||     ||||||              :::
Db        255 lySerProGlyProValGlyAlaProGlyProAlaGlyProSerGlyGlnProGlyGluA 275

Qy       2286 TAGGA------------------------------------------GGCTAGCTGTTAA 2269
                |||                                            |||      |
Db        275 rgGlyArgThrGlyProAlaGlySerGlnGlyAspArgGlyAlaAspGly------AlaT 293

Qy       2268 CCCTGAGCCTGGGTAATCCA-----------CCTGCAGAGTCCCCG------------ 2234
              ||   |||   |||   |||              ||||||   :::|||
Db        293 hrGlySerGlnGlyProProGlySerThrGlyProAlaGlyAlaProGlyMetProGlyI 313

Qy       2233 ----------CATTCCAGTGCATGGAGCCCTTCT------GGCCTCCCTGTATAAGTCC 2191
                        |||   ||||||           |||   |||
Db        313 leSerGlyAlaLysGlyAspAlaGlySerProGlyAlaArgGlySerPro---------- 329

Qy       2190 AGACTGAAACCCCCTTGGAAGGCCTCCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGG- 2132
                                |||||||||                       |||||||||
Db        330 ------------------GlyLeuGlnGlyAlaArg------------GlyGluArgG 339

Qy       2131 -------------AGAGGGACGCCCCAGCCCCCAGCTGTGCAGCTACGC---------- 2096
                            :::|||          ||||||   |||       |||
Db        339 lySerGluGlySerGlnGlyGlnThrGlyProProGlyValProGlyArgAspGlySerA 359

Qy       2095 ----------------------ACCTCAGCAGCACAGGGT--------------GGCA 2074
                                  |||   |||||||||                 |||:
Db        359 snGlyAlaLysGlySerAlaGlyProSerGlyAlaGlnGlyThrProGlyPheProGlyA 379

Qy       2073 GCAGA--------------------GAGCCACATTACTTTGGCAGCAACAGAAACTGGC 2035
              :::|||                        |||          ||||||      :::
Db        379 laArgGlyProProGlyProAlaGlySerProGlyProAlaGlySerLysGlyAspGlnG 399

Qy       2034 GGCCAGCCCGGCAGCCCCATGGGGCTAACAGGAGCGGGGAGCTGGGACCC---AGTGAGG 1978
```

```
Qy   2165 CCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCC-------------- 2121
                            :::   |||   ||||||:::|||  |||
Db    129 lyProProArgArg---AspGlyGlyAsnArgGlyAspGly-AlaProGluArgGlySer 147

Qy   2120 ------CCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGA 2067
                |||   |||   |||   :::      |||   |||
Db    148 GlyLeuProGlyProGlyCysGlySerProGlyLeuGlySer----------------- 161

Qy   2066 GCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGG------ 2013
                                              ||||||   |||   |||
Db    162 --------------------------------ArgSerProProAlaGlyTrpGlyGlu 170

Qy   2012 -----------------GGCTAACAGGAGCGGGGAGCT------GGGACCCAGTGAGGC 1977
                            |||   :::||||||   |||      |||   :::
Db    171 ProGlyGlyArgGlyProGlyAlaArgGluArgAlaAlaArgProGlyValArgValPro 190

Qy   1976 AGGCCCTCCACCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCT 1917
          ||||||      |||:::
Db    191 ArgProGlyGluProVal------------------------------------------ 196

Qy   1916 TGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCCTGCGGCAGACA 1857
                                              |||      |||:::         |||
Db    197 -----------------------------------ProProGlyGlyMetGlyGlyThr 204

Qy   1856 -----------------------------------CCATATAGGCAGTGACAGACTGGC 1833
                                              |||   |||
Db    205 GlyGlyThrGlyProArgSerGluGlyAlaGlyCysProAlaArg-------------- 219

Qy   1832 TGAGCTGGACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCAC 1773
                |||         |||      |||   |||||||||         |||   |||
Db    220 ------GlyAlaGlyProProAlaTrpGlyAlaGlyProProArgArgAspGly-GlyAs 237

Qy   1772 TATCCAGGATGGCGAGGTCCAGGCAGATGCC-----------CCGGCCCGGAACCACCC 1725
                ||||||      :::      :::                  |||          ||
Db    237 nArgGlyAspGlyAlaProGluArgGlySerGlyLeuProGlyProGlyCysGlySerPr 257

Qy   1724 TGGCCTCGGTGGGCTCACCCACCACCACACGTACGGAGACATCACAGGCAGAGGCCCCGC 1665
          |||||||||      :::            :::|||:::      ||||||||||||
Db    257 oGlyLeuGlySerArgSerProProAlaGlyTrpGlyGluProGlyGlyArgGlyProGl 277

Qy   1664 AGAGCGC-----------------GGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCAC 1623
             |||                  |||   |||         ||||||         |||    |
Db    277 yAlaArgGluArgAlaAlaArgProGlyValArg-ValProArgProGlyGluProValP 297

Qy   1622 CCACGTGTCCATTAGGGAAGGGAGCTCCAGGCTTAGGGCCTGGCAGGAAGCTGGTCATCA 1563
          ||          :::|||     |||      |||   ||||||         |||   |||
Db    297 roProGlyGlyMetGly---GlyThrGlyGlyThrGlyPro---Arg-SerGluGlyAla 314

Qy   1562 GGCTGTCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCT 1503
          |||||||||                     |||            ||||||      |||
Db    315 GlyCysProAla------ArgGlyAlaGlyProProAlaTrpGlyAlaGlyProProArg 332

Qy   1502 TCTCCCGGTGGT---AGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGG 1446
                ||||||   ||||||                        ||||||   ||||||
Db    333 ArgAspGlyGlyAsnArgGlyAspGlyAlaProGluArgGlySer---GlyLeu------ 349

Qy   1445 TGAACCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACA---GGCATG 1389
                                        |||   |||   ||||||:::   |||:::
Db    350 -----------------------------ProGlyProGlyCysGlySerProGlyLeu 359

Qy   1388 TGGCACCGGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGA 1329
                |||   |||
Db    360 GlySerArgSerPro--------------------------------------------- 364

Qy   1328 ATCGCTGCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGG---CGCACTGCA 1272
                   |||||||||                           |||
Db    365 ---------ProAlaGly------------------------TrpGlyGluProGly 372

Qy   1271 GGAACAGCCCCAGGCTGCCCATCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGC 1212
          |||      ||||||         ||||||
Db    373 GlyArgGlyProGlyAla---ArgGluArgAla------------------------- 382

Qy   1211 CCGGCTCAGCTCTGGGCACGCCCTGGTACA---GCCCCTCGCCCACGAAATCCGTGTAAA 1155
                         ||||||||||||      |||||||||      |||
Db    383 ---------------AlaArgProGlyValArgValProArgProGlyGluPro------ 395

Qy   1154 ACAGCGTGAAGGTCATGAGTGCCA--------------TCCAGCTGCACAGCTCAGCCA 1110
                            ||||||                  :::            |||
Db    396 -----------------ValProProGlyGlyMetGlyGlyThrGlyGlyThrGlyPro 409

Qy   1109 CGAAGAGCCGGCGCAGGGTGCGGGGCA------TGCGGCAGCACAGCTGGTGCAGCCGGG 1056
          |||      |||||||||||   |||         |||      |||      ||||||
Db    410 ArgSerGluGlyAlaGlyCysProAlaArgGlyAlaGlyProProAlaTrpGlyAlaGly 429

Qy   1055 ------------GAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGC--------------- 1023
                      :::   |||      |||   |||   |||
Db    430 ProProArgArgAspGlyGlyAsnArgGlyAspGlyAlaProGluArgGlySerGlyLeu 449

Qy   1022 ---GGCCCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCC--------- 975
             |||||||||      |||   |||         |||   |||   |||
Db    450 ProGlyProGlyCysGlySerProGlyLeuGlySerArgSerProProAlaGlyTrpGly 469

Qy    974 -----------------------------CTTCTGCTGGCTCGGTGGGGCCCAGCGCTG 945
                                         ||||||   |||         :::
Db    470 GluProGlyGlyArgGlyProGlyAlaArgGluArgAlaAlaArgProGlyValArgVal 489

Qy    944 CCT-----------CCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGA 897
          |||            |||   ||||||
Db    490 ProArgProGlyGluProValProProGly---------------------------- 499

Qy    896 GGGTGAGCAGGCCAAAGAGGCACTCCTCCT--------------GGGTGCCCAGGTAGG 852
          |||      |||      ||||||   |||                  |||||||||   |||
Db    500 GlyMetGlyGlyThrGlyGlyThrGlyProArgSerGluGlyAlaGlyCysProAlaArg 519

Qy    851 GGGCCAGGGCACTGGTGTCCCAGTCAATGGCAGGCAGGA------------------- 813
          |||   |||         |||      |||   ||||||
Db    520 GlyAlaGly---------ProProAlaTrpGlyAlaGlyProProArgArgAspGlyGly 536

Qy    812 ---GGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGC 756
                      |||:::|||:::      |||                        |||
Db    537 AsnArgGlyAspGlyAlaProGluArgGlySer---------------------Gly 548

Qy    755 GACAGTGGTCCGGGT---------CCCGGAAGAGGTCAGAGAGCAGGGCCTCCAGTGGAG 705
                |||||||||         ||||||   |||   |||:::   ||||||
Db    549 LeuProGlyProGlyCysGlySerProGlyLeuGlySerArgSer---ProProAlaGly 567

Qy    704 TGAAGCACACCTGGCCACAGAAGTCCAGCA-----------GCCCCACGCCCAGGATGA 657
                   ||||||   |||   |||               |||   |||||||||
Db    568 TrpGlyGluProGlyGlyArgGlyProGlyAlaArgGluArgAlaAlaArgProGlyVal 587

Qy    656 GCAGTGCCAGCTCCAGGG---GCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGG 600
             ||||||   ||||||               ||||||
Db    588 ArgValProArgProGlyGluProValProProGly--------------------- 599

Qy    599 CCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCAAGGACAGTGCCCAGATGAAGGGCC 540
                            |||      |||         |||   ||||||
Db    600 -----------------GlyMetGlyGlyThrGlyGlyThrGlyProArg--------- 610

Qy    539 GGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTAGGAGCGGGACACAGA 480
                                                 |||   |||||||||
Db    611 ------------------------------------SerGluGlyAlaGly------ 615

Qy    479 CCAGGCCCAGCACTGGACCAATGCCCAGCACCATGGTCATGAACTTCTCCTCTACCCCCA 420
                               |||||||||
Db    616 --------------------CysProAla---------------------------- 618

Qy    419 CTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCCAAACACACCTCCAGGCCAA 360
                                              |||            ||||||
Db    619 ------------------------------------ArgGlyAlaGlyProProAlaTrp 626

Qy    359 AGGTTAGCAGGTTGACCAGCAAGAGCTGGGCTTTCCGGTGCCGCAGCAGGCGGCTCACCC 300
                ||||||   |||   |||   |||      |||            |||:::|||
Db    627 Gly---AlaGlyProProArgArgAspGly------GlyAsnArgGlyAspGlyAlaPro 643

Qy    299 ACAGCCTCTGGACCATAGTGGGCCAGGCGGGTAGGGCTCAGGGGGCCGTTCAGGCACTCC 240
```

```
Db        278 euTyrPheAlaGluGluValProLeuMetAlaTyrGlnProHisHisLeuSerAspSer- 297

Qy        977 CTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTT---- 1032
                                  |||||||||||       |||        :::||||||
Db        298 -------------------AlaProLeuLeuAspAsnProGlnGlnIleGlyPheAspA 311

Qy       1033 -------------------------------------------------------CCGGAA- 1038
                                                                     ||||||
Db        311 snSerLysSerLysLeuAspMetSerAlaValAspAsnAlaThrGlyAsnAsnProGluS 331

Qy       1038 ------------------------------------------------------------ 1038

Db        331 erSerTyrGluIleAsnLysAsnAlaLysHisLeuThrProIleValGlnGluGlnAsnG 351

Qy       1039 -----------------CCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGC 1081
                               ||| ||||||:::|||      |||
Db        351 luSerPheSerAspGlyPro-GlyAlaValLeuValAsnLeuLeuThrSerLeuArgHis 370

Qy       1082 ATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATG 1141
              :::|||       :::       :::    :::            ||||||::::::
Db        371 LeuProProAlaMetHisSerValLeuLeuValMetAlaLeuSerTrpLeuSerTrpPhe 390

Qy       1142 ACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGA 1201
                 |||    ||||||    |||||::::::|||      :::|||    |||    |||:::
Db        391 ProPhePheLeuPheAspThrAspTrpMetGlyArgGluValTyrHisGlyAspProLys 410

Qy       1202 GCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTG 1261
                 :::                 |||   :::   ||||||   |||||||||   |||:::
Db        411 GlyAsp------GluSerAlaValLysAlaTyrAspAlaGlyValArgGluGlyAlaPhe 428

Qy       1262 GGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTG 1321
              ||||||   |||                      :::    |||   :::::::::    :::
Db        429 GlyLeuLeuLeuAsnSerValAspLeuGlyIleSerSerPheLeuIleGluProMetCys 448

Qy       1322 CAGCGATTCGGCACTCGAGCAGTCTAT---TTGGCCAGTGTGGCAGCTTTCCCTGTGGCT 1378
              ||||||   |||   |||   |||:::    :::::::::          |||
Db        449 GlnArgMetGlyAlaArgLeuValTrpAlaMetSerAsnPheIleValPheAlaCysMet 468

Qy       1379 GCCGGTGCCACATGCCTGTCC--------------------------------------- 1399
              ||||||         :::|||
Db        469 AlaGlyThrAlaIleIleSerLeuValSerValAsnGluTyrIleThrGluGlyIleGln 488

Qy       1400 CACAGTGTG-----------GCCGTGGTGACAGCTTCAGCC---------GCCCTCACC 1438
              |||::::::           |||:::      ||||||            ||||||
Db        489 HisAlaIleGlyGluAsnArgAlaIleLysIleAlaSerLeuValValPheAlaLeuLeu 508

Qy       1439 GGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGG 1498
              ||||||      |||               :::|||:::::::::    :::
Db        509 GlyPheProLeuSerIleThrTyrSerValProPheSerIleThrAla------------ 524

Qy       1499 GAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
                                        :::         ||||||||||||
Db        525 --------------------GluLeuThrAlaAspThrGlyGly 532
```

RESULT 9
O41972

```
ID   O41972      PRELIMINARY;      PRT;    706 AA.
AC   O41972;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   Hypothetical 66.4 kDa protein.
GN   GAMMAHV.M10B.
OS   murid herpesvirus 4.
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae.
OX   NCBI_TaxID=33708;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WUMS;
RX   MEDLINE=97366649; PubMed=9223479;
RA   Virgin H.W. IV, Latreille P., Wamsley P., Hallsworth K., Weck K.E.,
RA   Dal Canto A.J., Speck S.H.;
RT   "Complete sequence and genomic analysis of murine gammaherpesvirus
RT   68.";
RL   J. Virol. 71:5894-5904(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WUMS;
RA   Latreille P., Wamsley P., Waterston R.H.;
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U97553; AAB66421.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   706 AA;  66413 MW;  9AD7A4AFDEB8748B CRC64;

Alignment Scores:
Pred. No.:              1.84e-15      Length:         706
Score:                  365.50        Matches:        272
Percent Similarity:     29.39%        Conservative:   41
Best Local Similarity:  25.54%        Mismatches:     278
Query Match:            5.87%         Indels:         474
DB:                     12            Gaps:           61

US-09-759-143-110 (1-3410) x O41972 (1-706)

Qy       2939 GGAAGTGGGGGGAACCAGGCTGGGCCAAGAGAAGAGGGGTGGTTAGGGAAGCCGTTGAGA 2880
               |||:::||||||             |||||||||  ||||||      |||   |||   |||
Db          3 GlyThrGlyGly------ThrGlyProArgSerGluGly---AlaGlyCysProAlaArg 19

Qy       2879 CCTGAAGCCCCACCCTCTACCTTCCTTCAACACCCTAACCTTGGGTAACAGCATTTGGAA 2820
                           ||||||:::
Db         20 GlyAlaGlyProProAla------------------------------------------ 25

Qy       2819 TTATCATTTGGGATGAGTAGAATTTCCAAGGTCCTGGGTTAGGCATTTTGGGGGGCCAGA 2760
                                                                 |||||||||
Db         26 ------------------------------------------------TrpGlyAlaGly 29

Qy       2759 CCCCAGGAGAAGAAGATTCTGGCAATGATCAGCCCAATGACCAGCTATCTCAGGGGACCT 2700
               |||       :::
Db         30 ProProArgArg------------------------------------------------ 33

Qy       2699 GATTGTTGGGGATCCCCCACCCTACCCAAATATTAGACACCAACACAGAAAAGCTAGCAA 2640

Db         33 ------------------------------------------------------------ 33

Qy       2639 TGGATTCCCTTCTACTTTGTTAAATAAATAAGTTAAATATTTAAATGCCTGTGTCTCTGT 2580

Db         33 ------------------------------------------------------------ 33

Qy       2579 GATGGCAACAGAAGGACCAACAGGCCACATCCTGATAAAAGGTAAGAGGGGGGTGGATCA 2520
               ||||||    |||                                      |||
Db         34 AspGlyGlyAsnArg-------------------------------------GlyAspGly 41

Qy       2519 GCAAAAAGACAGTGCTGTGGGCTGAGGGGACCT------GGTTCTTGTGTGTTGCCCCTC 2466
               |||      :::      ||||||   ||||||      ||||||      |||
Db         42 AlaProGluArgGlySerGlyLeuProGlyProGlyCysGlySerProGlyLeuGlySer 61

Qy       2465 AGGACTCTTCCCCTACAAATAAGTCATATGTTCAAATCCCATGGAGGAGTGTTTCATCCT 2406
               |||:::   |||                        :::      ||||||
Db         62 ArgSerProProAlaGly------------TrpGlyGluProGlyGly------------ 73

Qy       2405 AGAAACTCCCATGCAAGAGCTACATTAAACGAAGCTGCAGGTTAAGGGGCTTAGAGATGG 2346
               |||         ||||||                ||||||        |||:::|||
Db         74 ArgGlyProGlyAlaArg---------GluArgAlaAlaArg-ProGlyValArgValPr 90

Qy       2345 GAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCTCAACT 2286
               :::||||||:::   |||                   |||       |||:::
Db         90 oArgProGlyGluProVal------------------ProProGlyGlyMetGlyGlyTh 104

Qy       2285 AGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCAGAGTCCCCGCATTCCAG 2226
               ||||||                                            |||    |||
Db        104 rGlyGly------------------------------------ThrGlyProArgSerGl 112

Qy       2225 TGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGCCT 2166
                              :::|||   |||                    |||   |||   |||
Db        112 uGly---------AlaGlyCysProAla-ArgGlyAlaGlyProProAlaTrpGlyAlaG 129
```

```
                    |||||||||       ||||||  |||   |||:::   ||||||
Db        577 LeuProGlyProGlyCysGlySerProGlyLeuGlySerArgSer---ProProAlaGly 595

Qy        704 TGAAGCACACCTGGCCACAGAAGTCCAGCA-----------GCCCCACGCCCAGGATGA 657
                         ||||||  |||  |||                |||  |||||||||
Db        596 TrpGlyGluProGlyGlyArgGlyProGlyAlaArgGluArgAlaAlaArgProGlyVal 615

Qy        656 GCAGTGCCAGCTCCAGGG---GCCTGGGATCCGGGCACAGCAGCCCTGCTAGCCAGCCGG 600
                 ||||||  ||||||          ||||||
Db        616 ArgValProArgProGlyGluProValProProGly---------------------- 627

Qy        599 CCCTTGGGATGAGAAAGAGGCTCAGCAGGATGCCCAAGGACAGTGCCCAGATGAAGGGCC 540
                                  |||     |||        |||  ||||||
Db        628 -----------------GlyMetGlyGlyThrGlyGlyThrGlyProArg--------- 638

Qy        539 GGCGGCGGCCATAGCGTCCACGCCAGTGGTCACTGGCTGAGCCTAGGAGCGGGACACAGA 480
                                                      |||  |||||||||
Db        639 -------------------------------------SerGluGlyAlaGly------ 643

Qy        479 CCAGGCCCAGCACTGGACCAATGCCCAGCACCATGGTCATGAACTTCTCCTCTACCCCCA 420
                                   |||||||||                     ||||||
Db        644 --------------------CysProAla-----------ArgGlyAlaGlyProPro 652

Qy        419 CTTCCAGCAGCAGAGGCGGCACATAGGTGATGCCTGCGGCCAAACACACCTCCAGGCCAA 360
                       |||                                  |||       |||
Db        653 AlaTrpGlyAlaGlyPro--------------------ProArgArgAspGlyGlyAsn 665

Qy        359 AGGTTAGCAGGTTGACCAGCAAGA---GCTGGGCTTTCCGGTGCCGCAGCAGGC------ 309
               |||     |||   |||   |||    :::||||||   |||          |||
Db        666 ArgGlyAspGlyAlaProGluArgGlySerGlyLeuProGlyProGlyCysGlySerPro 685

Qy        308 GGCTCACCCACAGCCTCTGGACCATAGTGGG---------CCAGGCGGGTAGGGCTCAGG 258
               |||     :::   |||   ||| :::|||          ||||||||   |||   ||
Db        686 GlyLeuGlySerArgSerProPro-AlaGlyTrpGlyGluProGlyGlyArgGlyProGl 705

Qy        257 GGGCCGTTCAGGCACTCCAGAACTGCTTCGTCTCGGCTCTGCTCCAGAAGCTGCGGCCTC 198
               |  |||                                       |||   ||||||
Db        705 yAlaArg---------------------------------------GluArgAlaAlaAr 712

Qy        197 TCCTCCTTGCTGCCGCCAACTGCCTAGGAATCAGCCAGGCGCCCATTTCTGCCAGCCCTT 138
                |||              :::::||||||       ||||||          :::|||
Db        712 gProGlyVal------ArgValProArg------ProGly-----------GluProVa 724

Qy        137 T--------------------------GGTGCCGGTCCAGCTTCTCAGCCCATGCTCAA 105
                                         |||   ||||||   |||:::
Db        724 lProProGlyGlyMetGlyGlyThrGlyGlyThrGlyProArgSerGlu---------Gl 741

Qy        104 CACCTGCTGC-----------TGTGGGGCACCTCAGTGGGGA---------------CA 72
                      |||             |||   |||   ||||||
Db        741 yAlaGlyCysProAlaArgGlyAlaGlyProProAlaTrpGlyAlaLysAlaArgSerSe 761

Qy         71 CGTCTCATCACTCAGATCCTGGCCGAGGCGCGCGGCT 35
                    :::   :::|||           ||||||   |||
Db        761 rThrAlaGluValArgProValAlaArgArgGlyAla 773

RESULT 8
Q9SQK5
ID   Q9SQK5      PRELIMINARY;      PRT;   612 AA.
AC   Q9SQK5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Putative sucrose transporter.
GN   VVSUC12.
OS   Vitis vinifera (Grape).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Vitaceae;
OC   Vitis.
OX   NCBI_TaxID=29760;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SHIRAZ; TISSUE=FRUIT;
RA   Davies C., Wolf T., Robinson S.P.;
RT   "Three putative sucrose transporters are differentially expressed in
RT   grapevine tissues.";
RL   Plant Sci. 147:93-100(1999).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR   EMBL; AF021809; AAF08330.1; -.
DR   TIGRFAMs; TIGR01301; GPH_sucrose; 1.
KW   Transmembrane.
SQ   SEQUENCE   612 AA;  65711 MW;  8FC3FBD6AB439078 CRC64;

Alignment Scores:
Pred. No.:              1.55e-15      Length:         612
Score:                  366.50        Matches:        142
Percent Similarity:     40.60%        Conservative:   76
Best Local Similarity:  26.44%        Mismatches:     193
Query Match:            5.71%         Indels:         127
DB:                     10            Gaps:           19

US-09-759-143-110 (1-3410) x Q9SQK5 (1-612)

Qy        219 GAGCCGAGACGAAGC---------AGTTCTGGAGTGCCTGAACGGCCCCCTGAGCCCTAC 269
              |||||||||                :::|||  |||||| ||||||         |||::
Db         29 GluProArgHisGlyAlaAspLeuAsnSerArgValPro-AsnGlyThrSerAspProSe 48

Qy        270 CCGCCTGGCCCACTATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAG 329
               :                    |||   ::::::     ||||||
Db         48 rSer------------SerProSerSerIleProHisProProLys-HisGlyGlyLeuA 64

Qy        330 CCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCA 389
                    ||::::|||      ::::::      |||::::::        |||     :
Db         64 rgThrLeuIleLeuSerCysMetIleAlaAlaGlyValGlnPheGlyTrpAlaLeuGlnL 84

Qy        390 TCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGG 449
               ::::   :::   |||   :::      :::|||:::|||      |||   :::   :
Db         84 euSerLeuLeuThrProTyrIleGlnThrLeuGlyIleGluHisAlaPheSerSerPheI 104

Qy        450 TGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTG 509
               ::          ||||||:::  |||||||||    |||   :::|||     ||||
Db        104 leTrpLeuCysGlyProIleThrGlyLeuValValGlnProCysValGlyIleTrpSerA 124

Qy        510 ACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCA 569
               ||            :::|||||||||||||||||||||||   |||   ||||||   |
Db        124 spLysCysSerSerLysTyrGlyArgArgArgProPheIleLeuAlaGlySerLeuMetI 144

Qy        570 TCCTGCTGAGCCTCTTTCTCATC---------CCAAGGGCCGGCTGGCTAGCAGGGCTGC 620
               ||   :::::::::   :::|||                    |||:::|||   |||
Db        144 leSerValAlaValThrIleIleGlyPheSerAlaAspIleGlyTyrLeuLeuGlyAspT 164

Qy        621 TGTGCCCGGATCCCAGGCCCCTGGAG-----------------CTGGCACTGCTCATCC 662
                         |||   |||     :::                     :::   :::|
Db        164 hrAsnMetAspCysArgLysPheLysGlyThrArgThrTrpAlaAlaIleIlePheValL 184

Qy        663 TGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGC 722
              |||||      :::||||||                          |||    |||||||
Db        184 euGlyPheTrpMetLeuAspLeuAlaAsnAsnThrValGlnGlyProAlaArgAlaLeuL 204

Qy        723 TCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCA 782
              ||:::||||||              ||||||          |||   :::::::::
Db        204 euAlaAspLeu---SerGlyProAspGlnArgAsnSerAlaAsnAlaIlePheCysSerT 223

Qy        783 TGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC-----------T 830
                 :::::::::|||       ||||||:::          |||               |
Db        223 rpMetAlaValGlyAsnIleLeuGlyPheSerAlaGlyAlaSerGlyHisTrpHisArgT 243

Qy        831 GGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGC------------ 877
              ||                   |||:::|||       :::      |||
Db        243 rp---------------PheProPheLeuLeuAsnLysAlaCysCysGluAlaCysGlyA 258

Qy        878 --------CTCTTTGGCCTGCTCACCCTCATCTTCCTCACC---TGCGTAGCAGCCACAC 926
                              |||:::  ::::::|||||||||   |||         ||||
Db        258 snLeuLysAlaAlaPheLeuIleAlaValValPheLeuThrLeuCysThrLeuValThrL 278

Qy        927 TGCTGGTGGCTGAGGAGGCAGCGCTGGG----------CCCCACCGAGCCAGCAGAAGGG 976
               ||     |||||||||      |||                  ||||||      |||
```

```
Qy       2639 TGGATTCCCTTCTACTTTGTTAAATAAATAAGTTAAATATTTAAATGCCTGTGTCTCTGT 2580

Db         61 ------------------------------------------------------------ 61

Qy       2579 GATGGCAACAGAAGGACCAACAGGCCACATCCTGATAAAAGGTAAGAGGGGGGTGGATCA 2520
              ||||||     |||                                  |||
Db         62 AspGlyGlyAsnArg----------------------------------GlyAspGly 69

Qy       2519 GCAAAAAGACAGTGCTGTGGGCTGAGGGGACCT------GGTTCTTGTGTGTTGCCCCTC 2466
              |||    :::    ||||||  ||||||        ||||||     |||
Db         70 AlaProGluArgGlySerGlyLeuProGlyProGlyCysGlySerProGlyLeuGlySer 89

Qy       2465 AGGACTCTTCCCCTACAAATAAGTCATATGTTCAAATCCCATGGAGGAGTGTTTCATCCT 2406
              |||:::  |||                 :::      ||||||
Db         90 ArgSerProProAlaGly-----------TrpGlyGluProGlyGly----------- 101

Qy       2405 AGAAACTCCCATGCAAGAGCTACATTAAACGAAGCTGCAGGTTAAGGGGCTTAGAGATGG 2346
              |||        ||||||                ||||||     |||:::|||
Db        102 ArgGlyProGlyAlaArg---------GluArgAlaAlaArg-ProGlyValArgValPr 118

Qy       2345 GAAACCAGGTGACTGAGTTTATTCAGCTCCCAAAAACCCTTCTCTAGGTGTGTCTCAACT 2286
               :::|||||||:::   |||                 |||       |||:::
Db        118 oArgProGlyGluProVal------------------ProProGlyGlyMetGlyGlyTh 132

Qy       2285 AGGAGGCTAGCTGTTAACCCTGAGCCTGGGTAATCCACCTGCAGAGTCCCCGCATTCCAG 2226
               ||||||                                  |||    |||
Db        132 rGlyGly-----------------------------------ThrGlyProArgSerGl 140

Qy       2225 TGCATGGAGCCCTTCTGGCCTCCCTGTATAAGTCCAGACTGAAACCCCCTTGGAAGGCCT 2166
                          :::|||   |||     :::                |||   |||   |||
Db        140 uGly---------AlaGlyCysProAla-ArgGlyAlaGlyProProAlaTrpGlyAlaG 157

Qy       2165 CCAGTCAGGCAGCCCTAGAGACTGGGGAGAGAGGAGAGGGACGCC--------------- 2121
                             :::   |||    ||||||:::|||  |||
Db        157 lyProProArgArg---AspGlyGlyAsnArgGlyAspGly-AlaProGluArgGlySer 175

Qy       2120 ------CCAGCCCCCAGCTGTGCAGCTACGCACCTCAGCAGCACAGGGTGGCAGCAGAGA 2067
                 |||    |||   |||    :::      |||   |||
Db        176 GlyLeuProGlyProGlyCysGlySerProGlyLeuGlySer----------------- 189

Qy       2066 GCCACATTACTTTGGCAGCAACAGAAACTGGCGGCCAGCCCGGCAGCCCCATGG------ 2013
                                              ||||||    |||    |||
Db        190 ---------------------------------ArgSerProProAlaGlyTrpGlyGlu 198

Qy       2012 -----------------GGCTAACAGGAGCGGGGAGCT------GGGACCCAGTGAGGC 1977
                                 |||    :::||||||   |||         |||    :::
Db        199 ProGlyGlyArgGlyProGlyAlaArgGluArgAlaAlaArgProGlyValArgValPro 218

Qy       1976 AGGCCCTCCACCCCAATGTGCTGGAAGTTTTCTACGCTGAGTATTTGGCCAAGTCGCTCT 1917
              ||||||     |||:::
Db        219 ArgProGlyGluProVal------------------------------------------ 224

Qy       1916 TGTCAAATACTACCTGTGTAGCAAAGTAAATGGCGACCAGACCCAGGCCTGCGGCAGACA 1857
                                                   |||     |||:::      |||
Db        225 ----------------------------------ProProGlyGlyMetGlyGlyThr 232

Qy       1856 ----------------------------------CCATATAGGCAGTGACAGACTGGC 1833
                                                  |||    |||
Db        233 GlyGlyThrGlyProArgSerGluGlyAlaGlyCysProAlaArg-------------- 247

Qy       1832 TGAGCTGGACAATGGAGCCCATAAACAGGGATGGGGCCACCTGGGACAGCAGGAAGGCAC 1773
                    |||       |||   |||      |||||||||         |||  |||
Db        248 ------GlyAlaGlyProProAlaTrpGlyAlaGlyProProArgArgAspGly-GlyAs 265

Qy       1772 TATCCAGGATGGCGAGGTCCAGGCAGATGCC------------CCGGCCCGGAACCACCC 1725
                   ||||||        :::          :::          |||            ||
Db        265 nArgGlyAspGlyAlaProGluArgGlySerGlyLeuProGlyProGlyCysGlySerPr 285

Qy       1724 TGGCCTCGGTGGGCTCACCCACCACCACACGTACGGAGACATCACAGGCAGAGGCCCCGC 1665
              ||||||||||          :::               :::|||:::       |||||||||||
Db        285 oGlyLeuGlySerArgSerProProAlaGlyTrpGlyGluProGlyGlyArgGlyProGl 305

Qy       1664 AGAGCGC-----------------GGGTGGAGGTGGGAGCAGGCCACTGCCTCCAGCAC 1623
                 |||             |||    |||        ||||||       |||   |
Db        305 yAlaArgGluArgAlaAlaArgProGlyValArg-ValProArgProGlyGluProValP 325

Qy       1622 CCACGTGTCCATTAGGGAAGGGAGCTCCAGGCTTAGGGCCTGGCAGGAAGCTGGTCATCA 1563
              ||           :::|||   |||        |||   ||||||   |||  |||
Db        325 roProGlyGlyMetGly---GlyThrGlyGlyThrGlyPro---Arg-SerGluGlyAla 342

Qy       1562 GGCTGTCCTCACTGCTAGCACCTCCAGTGTCCCCTCGGTATTTGGGCAGGAACACCTGCT 1503
              |||||||||                             |||          ||||||   |||
Db        343 GlyCysProAla------ArgGlyAlaGlyProProAlaTrpGlyAlaGlyProProArg 360

Qy       1502 TCTCCCGGTGGT---AGAGGGAGGCCAGTGTGTAGGGCAGGATCTGCAGGGCTGAGAAGG 1446
                    ||||||   ||||||                     ||||||   ||||||
Db        361 ArgAspGlyGlyAsnArgGlyAspGlyAlaProGluArgGlySer---GlyLeu------ 377

Qy       1445 TGAACCCGGTGAGGGCGGCTGAAGCTGTCACCACGGCCACACTGTGGGACA---GGCATG 1389
                                             |||   |||   ||||||:::   |||:::
Db        378 ------------------------------ProGlyProGlyCysGlySerProGlyLeu 387

Qy       1388 TGGCACCGGCAGCCACAGGGAAAGCTGCCACACTGGCCAAATAGACTGCTCGAGTGCCGA 1329
                    |||   |||
Db        388 GlySerArgSerPro--------------------------------------------- 392

Qy       1328 ATCGCTGCACCAGCCGGTCCATGACCAGAGAGAAGACCAGGGAGATGG---CGCACTGCA 1272
                       |||||||||                            |||
Db        393 ---------ProAlaGly---------------------------TrpGlyGluProGly 400

Qy       1271 GGAACAGCCCCAGGCTGCCCATCCGAACGCCTTCATCATAGTGTCTCCGGGCCTCGGTGC 1212
              |||      ||||||        ||||||
Db        401 GlyArgGlyProGlyAla---ArgGluArgAla--------------------------- 410

Qy       1211 CCGGCTCAGCTCTGGGCACGCCCTGGTACA---GCCCCTCGCCCACGAAATCCGTGTAAA 1155
                               ||||||||||||      |||||||||       |||
Db        411 ---------------AlaArgProGlyValArgValProArgProGlyGluPro------ 423

Qy       1154 ACAGCGTGAAGGTCATGAGTGCCA---------------TCCAGCTGCACAGCTCAGCCA 1110
                                ||||||                     :::          |||
Db        424 -----------------ValProProGlyGlyMetGlyGlyThrGlyGlyThrGlyPro 437

Qy       1109 CGAAGAGCCGGCGCAGGGTGCGGGGCA------TGCGGCAGCACAGCTGGTGCAGCCGGG 1056
              |||      ||||||||||||   |||          |||        |||      ||||||
Db        438 ArgSerGluGlyAlaGlyCysProAlaArgGlyAlaGlyProProAlaTrpGlyAlaGly 457

Qy       1055 ------------GAAGCAGGGCGCCCAGGTTCCGGAAAGCCAAGC-------------- 1023
                          :::   |||       |||   |||   |||
Db        458 ProProArgArgAspGlyGlyAsnArgGlyAspGlyAlaProGluArgGlySerGlyLeu 477

Qy       1022 ---GGGCCCGGCATGGACAGCAGTGGGGCGACAAGGAGGGGGCCGACAGCC--------- 975
                 |||||||||   |||   |||          |||    |||    |||
Db        478 ProGlyProGlyCysGlySerProGlyLeuGlySerArgSerProProAlaGlyTrpGly 497

Qy        974 ----------------------------CTTCTGCTGGCTCGGTGGGGCCCAGCGCTG 945
                                                ||||||    |||         :::
Db        498 GluProGlyGlyArgGlyProGlyAlaArgGluArgAlaAlaArgProGlyValArgVal 517

Qy        944 CCT-----------CCTCAGCCACCAGCAGTGTGGCTGCTACGCAGGTGAGGAAGATGA 897
              |||              |||   ||||||
Db        518 ProArgProGlyGluProValProProGly----------------------------- 527

Qy        896 GGGTGAGCAGGCCAAAGAGGCACTCCTCCT--------------GGGTGCCCAGGTAGG 852
              |||      |||      ||||||   |||               |||||||||   |||
Db        528 GlyMetGlyGlyThrGlyGlyThrGlyProArgSerGluGlyAlaGlyCysProAlaArg 547

Qy        851 GGGCCAGGGCACTGGTGTCCCAGTCAATGGCAGGCAGGA-------------------- 813
              |||   |||         |||      |||      |||   ||||||
Db        548 GlyAlaGly---------ProProAlaTrpGlyAlaGlyProProArgArgAspGlyGly 564

Qy        812 ---GGTAGCCCAGGCAGCCCCCAAGACTGATCATGAAGGCATAGACAGAGTAGGCCTGGC 756
                         |||:::|||:::            |||                       |||
Db        565 AsnArgGlyAspGlyAlaProGluArgGlySer-----------------------Gly 576

Qy        755 GACAGTGGTCCGGGT---------CCCGGAAGAGGTCAGAGAGCAGGGCCTCCAGTGGAG 705
```

```
Db        219 AlaCysGlyTyrLeuValGlyAlaMetAspTrpGlyHisSerValLeuGlyArgLeuLeu 238

Qy        860 GGCACCCAGGAGGAGTGCCTCTTT--------------------GGCCTGCTCACCCTC 898
              |||:::::   :::   ::::::                      |||:::      :::
Db        239 GlySerGluTyrGlnValIleTyrPhePheSerAlaLeuThrTrpGlyValPheLeuIle 258

Qy        899 ATCTTCCTCACCTGCGTA------------------------------------------ 916
              :::   |||     :::
Db        259 ValHisLeuPheSerIleProGluLysProLeuAlaLysValProSerGluSerSerAla 278

Qy        917 ---GCAGCCACACTGCTGGTGGCTGAGGAGGCA-----------GCGCTGGGC------ 955
                 :::|||     |||:::            :::          |||||||||
Db        279 SerSerAlaLeuArgLeuLeuGlyProHisSerAsnGlyTyrGlyAlaLeuGlyLysGlu 298

Qy        956 CCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGC 1015
              |||      |||            |||:::|||   :::   |||
Db        299 ProValSerProValIleProThrSerSerProGluIleArgProArgSerTyrSerAla 318

Qy       1015 ------------------------------------------------------------ 1015

Db        319 LeuGlyGluArgProArgSerPheSerAlaLeuGlyGluAlaAsnSerValThrSerSer 338

Qy       1016 -----------------------CGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTG 1051
                                     :::   ::::::   |||||||:::
Db        339 AlaLysGlnProIleLysGluAspGlnLysLysMetThrPheArgSer----------- 354

Qy       1052 CTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTG 1111
                       |||   :::            ||||||      |||   |||   :::
Db        355 ---------LeuMetLysAlaIlePheAsnMetProAsnHisTyrArgPheLeuCysIle 371

Qy       1112 GCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTG 1171
              :::   |||      |||   |||   :::            ||||||:::||||||||::::
Db        372 SerHisLeuLeuGlyTrpAlaAlaPheLeuCysAsnMetLeuPhePheThrAspPheMet 391

Qy       1172 GGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACAC 1231
              |||:::   :::|||:::|||   |||   ||||||         :::   |||
Db        392 GlyGlnIleValTyrArgGlyAsnProTyrAlaGluHisAsnSerThrAlaTyrIleThr 411

Qy       1232 TATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTG 1291
              |||:::   ||||||   :::|||      ||||||   :::            |||
Db        412 TyrGluArgGlyValGluValGlyCysTrpGlyLeuCysIleAsnAlaValSerSerAla 431

Qy       1292 GTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTG 1351
              ::::::|||   |||            |||:::         |||   :::   :::|||
Db        432 LeuTyrSerTyrValGlnArgPheLeuLeuProTyrIleGlyLeuLysGlyLeuTyrPhe 451

Qy       1352 GCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCC 1411
                           |||   :::                     |||         |||
Db        452 MetGlyTyrPheValPheGlyMetGlyThrSerLeuIleGlyLeuPheProGluValIle 471

Qy       1412 GTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCC 1471
                                 ::::::   |||         |||   |||      :::|||
Db        472 AlaThrLeuIleLeuCysSerValPheGlyValMetSerSerThrLeuTyrThrIlePro 491

Qy       1472 TACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGG 1531
              :::   |||   :::   |||   ||||||::::::         :::   |||      |||
Db        492 PheAsnLeuIleAlaGluTyrGlnArgGluGluGluGluGlnVal---LysLeuGluGly 510

Qy       1532 GACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAG 1591

Db        510 ------------------------------------------------------------ 510

Qy       1592 CCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCT 1651
                                   |||:::            |||:::|||:::
Db        511 --------------------GlyAsnGluSerProArgGlyThrGlyMet--------- 520

Qy       1652 CCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAG 1711
                       |||   |||         |||   |||   :::
Db        521 ---------AspCysAlaAlaLeuThrCysMetValGlnLeu------------------ 531

Qy       1712 CCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGAT 1771
                       |||:::::::::   |||   |||
Db        532 ---------AlaGlnIleIleValGlyAlaGly--------------------------- 539

Qy       1772 AGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTC 1831
                                                    :::|||:::::::|||
Db        540 ------------------------------------LeuGlyAlaLeuValAsnAla 546

Qy       1832 AGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTAC 1891
              :::   ||||||            ::::::||||||:::   :::   |||:::
Db        547 AlaGlySerValIleValValValLeuSerAlaSerSerIleSerLeuIleGlyCysIle 566

Qy       1892 TTT 1894
              |||
Db        567 Phe 567
```

```
RESULT 7
O41971
ID   O41971      PRELIMINARY;      PRT;   774 AA.
AC   O41971;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Hypothetical 73.2 kDa protein.
GN   GAMMAHV.M10A.
OS   murid herpesvirus 4.
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae.
OX   NCBI_TaxID=33708;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WUMS;
RX   MEDLINE=97366649; PubMed=9223479;
RA   Virgin H.W. IV, Latreille P., Wamsley P., Hallsworth K., Weck K.E.,
RA   Dal Canto A.J., Speck S.H.;
RT   "Complete sequence and genomic analysis of murine gammaherpesvirus
RT   68.";
RL   J. Virol. 71:5894-5904(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WUMS;
RA   Latreille P., Wamsley P., Waterston R.H.;
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U97553; AAB66420.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   774 AA;  73164 MW;  D1B37F264C92C42C CRC64;

Alignment Scores:
Pred. No.:               1.5e-16      Length:          774
Score:                   382.50       Matches:         285
Percent Similarity:      29.90%       Conservative:    43
Best Local Similarity:   25.98%       Mismatches:      291
Query Match:             6.14%        Indels:          478
DB:                      12           Gaps:            67

US-09-759-143-110 (1-3410) x O41971 (1-774)

Qy       2939 GGAAGTGGGGGGAACCAGGCTGGGCCAAGAGAAGAGGGGTGGTTAGGGAAGCCGTTGAGA 2880
                 |||:::||||||            |||||||||   ||||||      |||   |||   |||
Db         31 GlyThrGlyGly------ThrGlyProArgSerGluGly---AlaGlyCysProAlaArg 47

Qy       2879 CCTGAAGCCCCACCCTCTACCTTCCTTCAACACCCTAACCTTGGGTAACAGCATTTGGAA 2820
                       ||||||:::
Db         48 GlyAlaGlyProProAla------------------------------------------ 53

Qy       2819 TTATCATTTGGGATGAGTAGAATTTCCAAGGTCCTGGGTTAGGCATTTTGGGGGGCCAGA 2760
                                                            |||||||||
Db         54 ------------------------------------------------TrpGlyAlaGly 57

Qy       2759 CCCCAGGAGAAGAAGATTCTGGCAATGATCAGCCCAATGACCAGCTATCTCAGGGGACCT 2700
                 |||      :::
Db         58 ProProArgArg------------------------------------------------ 61

Qy       2699 GATTGTTGGGGATCCCCCACCCTACCCAAATATTAGACACCAACACAGAAAAGCTAGCAA 2640

Db         61 ------------------------------------------------------------ 61
```

```
Qy       1700 GTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCG---------GGCCGGGGCATCTGC 1750
                                             :::||||||         |||||||:::
Db        533 -----------------------GluIleValProLeuLysGlnAlaArgGlyLeuGly 544

Qy       1751 CTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTT 1810
                 |||:::||||||:::    |||          ::::::|||:::       ||||||
Db        545 ThrAspValAlaIleIleSerSerMetValPheIleAlaGlnLeuIleValSerLeuSer 564

Qy       1811 ATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGC 1870
              :::|||   :::|||                :::      |||   :::   ::::::
Db        565 ValGlyProLeuValSerTrpMetAspThrThrCysAlaValLeuTyrAlaSerThrPhe 584

Qy       1871 CTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTT 1912
              |||      :::|||         |||      |||:::::
Db        585 LeuSerPheLeuAlaAlaIleAlaAlaMetPheValLeuTyr 598
```

RESULT 5
Q8R1I0

```
ID   Q8R1I0      PRELIMINARY;      PRT;   112 AA.
AC   Q8R1I0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical 11.4 kDa protein (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=COLON;
RA   Strausberg R.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC024519; AAH24519.1; -.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   112 AA;  11393 MW;  B9D9DE79FEF09FB7 CRC64;

Alignment Scores:
Pred. No.:              4.93e-25      Length:        112
Score:                  512.00        Matches:       102
Percent Similarity:     95.50%        Conservative:  4
Best Local Similarity:  91.89%        Mismatches:    5
Query Match:            7.98%         Indels:        0
DB:                     11            Gaps:          0

US-09-759-143-110 (1-3410) x Q8R1I0 (1-112)

Qy       1607 CCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGC 1666
              ||||||||||||||||||:::|||   ||||||:::|||   ||||||||||||||||||
Db          1 ProAsnGlyHisValGlySerGlySerSerGlyIleLeuAlaProProProAlaLeuCys 20

Qy       1667 GGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGG 1726
              ||||||||||||||||||||||:::|||||||||||||||||||||   |||||||||
Db         21 GlyAlaSerAlaCysAspValSerMetArgValValValGlyGluProProGluAlaArg 40

Qy       1727 GTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTG 1786
              ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ValValThrGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeuLeu 60

Qy       1787 TCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACT 1846
              ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db         61 SerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerHisSerValThr 80

Qy       1847 GCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTA 1906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGlnVal 100

Qy       1907 GTATTTGACAAGAGCGACTTGGCCAAATACTCA 1939
              ||||||||||||:::||||||||||||||||||
Db        101 ValPheAspLysAsnAspLeuAlaLysTyrSer 111
```

RESULT 6
Q90Z74

```
ID   Q90Z74      PRELIMINARY;      PRT;   576 AA.
AC   Q90Z74;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Membrane-associated transporter protein B.
OS   Oryzias latipes (Medaka fish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HNI;
RX   MEDLINE=21372467; PubMed=11479596;
RA   Fukamachi S., Shimada A., Shima A.;
RT   "Mutations in the gene encoding B, a novel transporter protein, reduce
RT   melanin content in medaka.";
RL   Nat. Genet. 28:381-385(2001).
DR   EMBL; AF332510; AAK77024.1; -.
SQ   SEQUENCE   576 AA;  63411 MW;  13A5C9828ABEE894 CRC64;

Alignment Scores:
Pred. No.:              2.62e-21      Length:        576
Score:                  456.00        Matches:       144
Percent Similarity:     39.77%        Conservative:  95
Best Local Similarity:  23.96%        Mismatches:    198
Query Match:            7.11%         Indels:        164
DB:                     13            Gaps:          17

US-09-759-143-110 (1-3410) x Q90Z74 (1-576)

Qy        317 CGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGT 376
              |||      :::   :::|||:::|||   :::::::::   ||||||   |||   |||
Db         56 ArgArgSerArgGlyArgLeuIleLeuHisSerMetValMetPheGlyArgGluPheCys 75

Qy        377 TTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAG 436
                 |||               :::|||   |||:::||||||   ||||||:::
Db         76 TyrAlaValGluAlaAlaPheValThrProValLeuLeuSerValGlyLeuProArgSer 95

Qy        437 TTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTA 496
                 ::::::|||      |||   |||:::||||||   :::      |||::::::
Db         96 LeuTyrSerLeuValTrpLeuIleSerProIleLeuGlyPheLeuLeuGlnProIleIle 115

Qy        497 GGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCA 556
              |||||||||||||||:::   |||         :::||||||||||||||||:::|||
Db        116 GlySerAlaSerAspTyrCysArgSerSerTrpGlyArgArgArgProTyrIle------ 133

Qy        557 CTGTCCTTGGGCATCCTG------------CTGAGCCTCTTTCTC--------------- 589
              |||   ||||||||||||            ||||||:::||||||
Db        134 LeuValLeuGlyIleLeuMetLeuValGlyLeuSerMetPheLeuAsnGlyAspAlaVal 153

Qy        590 -----------------------ATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGC 625
                                            |||:::   |||
Db        154 ValSerGluLeuValSerAspArgSerSerArgSerThrTrp------------------ 167

Qy        626 CCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTC 685
                                 :::   :::::::::   ||||||   |||   ||||||
Db        168 -----------------AlaIleValValValMetPheGlyValValLeuPheAspPhe 181

Qy        686 TGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTC------TTCCGG 739
                                 |||::::::|||   |||   |||:::       ::::::
Db        182 AlaAlaAspPheIleAspGlyProIleLysAlaTyrLeuPheAspValCysSerTyrGln 201

Qy        740 GACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGC 799
              |||   :::         :::            :::|||            |||||||||
Db        202 AspLysGlu---------ArgGlyLeuHisTyrHisAlaLeuPheThrGlyLeuGlyGly 218

Qy        800 TGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTG 859
                 |||||||||:::   |||:::||||||         |||   |||            |||
```

```
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003552; AAF50310.1; -.
DR   FlyBase; FBgn0035968; CG4484.
SQ   SEQUENCE   599 AA;  66057 MW;  C5381D334CFBF2EB CRC64;

Alignment Scores:
Pred. No.:              2.13e-25      Length:         599
Score:                  519.50        Matches:        150
Percent Similarity:     39.43%        Conservative:   100
Best Local Similarity:  23.66%        Mismatches:     201
Query Match:            8.09%         Indels:         183
DB:                     5             Gaps:           14

US-09-759-143-110 (1-3410) x Q9VSV1 (1-599)

Qy        317 CGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGT 376
              |||      :::   ::::::      ::::::   :::            :::|||
Db         46 ArgLysThrArgPheGluMetPheArgLeuSerAlaIleAlaMetAlaIleGluPheAla 65

Qy        377 TTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGTAGAGGAGAAG 436
                 ||||||      ::::::|||   |||:::||||||:::::|||||||:::   |||
Db         66 TyrAlaAlaGluThrSerPheValSerProIleLeuLeuGlnIleGlyValAspHisLys 85

Qy        437 TTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTA 496
                 |||:::|||      |||:::   |||:::::::|||         |||||||||
Db         86 HisMetSerMetThrTrpGlyLeuSerProLeuIleGlyPhePheMetSerProLeuLeu 105

Qy        497 GGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCA 556
              ||||||   ||||||      :::   |||:::|||||||||||||||   |||
Db        106 GlySerIleSerAspArgCysLysLeuArgTrpGlyArgArgArgProIleIleSerIle 125

Qy        557 CTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCA--------------------- 595
              ||||||   ||||||:::      |||   |||:::|||
Db        126 LeuSerPheGlyIleMetCysGlyLeuIleLeuValProTyrGlyLysAspLeuGlyLeu 145

Qy        596 ---------------------------------------------------------AGG 598

Db        146 LeuLeuGlyAspAlaGlyTyrThrTyrAlaGluSerAlaLeuAsnPheThrSerSerSer 165

Qy        599 GCCGGCTGGCTAGCAGGGCTGCTGTGC------------------CCGGATCCCAGGCCC 640
                 |||   :::|||   |||:::                  |||
Db        166 GlyGlySerValAlaAlaLeuValSerGlyGluAlaThrThrGlyProSerAlaSerAsp 185

Qy        641 CTGGAGCTGGCA------CTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 694
                 :::   |||      |||   |||||||||:::   |||||||||||
Db        186 TyrLysPheAlaValIleLeuThrIleLeuGlyMetValLeuLeuAspPheAspAlaAsp 205

Qy        695 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 754
                 |||   ||||||            |||   |||:::         |||:::
Db        206 ThrCysGlnThrProAlaArgThrTyrLeuLeuAspMetCys---ValProGluGluGln 224

Qy        755 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 814
                 :::|||   :::::::::|||                  ||||||   :::||||||
Db        225 ProLysAlaMetThrMetPheAlaLeuPheAlaGlyPheGlyGlyThrIleGlyTyrAla 244

Qy        815 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 874
              :::      :::||||||:::|||:::   :::         ::::::|||
Db        245 IleGlyGlyValAspTrpGluThrThrHisIleGlySerPheMetGlyGlyAsnIlePro 264

Qy        875 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTG------ 928
                 :::|||   |||:::|||:::||||||         |||         |||:::
Db        265 ThrValPheThrLeuValThrIleIlePheAlaValCysTyrLeuIleThrValThrThr 284

Qy        929 --------------------CTGGTGGCTGAGGAGGCAGCGCTGGGCCCC--------- 958
                                 |||:::   ::::::      |||   |||
Db        285 PheArgGluIleProLeuProLeuIleGluGlnAspGluLeuLeuArgProLeuSerGlu 304

Qy        958 ------------------------------------------------------------ 958

Db        305 GlnAlaIleLysLysGluLeuLysLysLysAsnAsnThrIleTyrTyrIleGlnGluThr 324

Qy        959 ------------------------------------------ACCGAGCCAGCAGAAGGG 976
                                                           |||      :::|||
Db        325 ThrGlnLeuGluLeuGlnMetAlaSerAspAspProLysArgLeuGluAlaLeuGlnGly 344

Qy        977 CTGTCGGCCCCCTCCTTGTCGCCC------------------------------------ 1000
                       ||||||
Db        345 SerTyrGlnAsnGlyTyrSerProAlaValGluLysGlnGlyLysSerGlnAspLeuGlu 364

Qy       1001 ------------CACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCC 1048
                          :::      |||         :::         :::::::::
Db        365 ThrGlnSerAspTyrAspAlaProValSerLeuLysAlaTyrLeuLysSerIlePheIle 384

Qy       1049 CTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTC 1108
                                               ||||||   ::::::|||   |||
Db        385 ---------------------------------MetProTyrSerMetArgMetLeuAla 393

Qy       1109 GTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTC 1168
              :::      |||   ||||||      :::|||:::   |||:::::::|||||||||
Db        394 LeuThrAsnLeuPheCysTrpMetGlyHisValThrTyrCysLeuTyrPheThrAspPhe 413

Qy       1169 GTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGA 1228
              |||||||||   ::::::   |||   |||   |||   |||   ::::::|||
Db        414 ValGlyGluAlaValPheHisGlyAspProThrAlaAlaProAsnSerGluAlaAlaLeu 433

Qy       1229 CACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCC 1288
              :::|||:::   |||||||||   |||      |||:::   :::
Db        434 AsnTyrGluAlaGlyValArgPheGlyCysTrpGlyMetAlaIleTyrAlaPheSerCys 453

Qy       1289 CTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTAT 1348
                 ::::::||||||   :::   :::|||::::::   |||||||||:::|||||||||
Db        454 SerIleTyrSerLeuSerValThrLysLeuMetLysTrpPheGlyThrLysAlaValTyr 473

Qy       1349 TTGGCCAGTGTGGCAGCTTTCCCTGTG---------GCTGCCGGTGCCACATGCCTGTCC 1399
              ::::::   :::      :::   :::                  |||
Db        474 IleSerGlyMetIleTyrTyrGlyIleGlyMetLeuValLeuGlyLeuTrpProThrLys 493

Qy       1400 CACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTG 1459
                 |||   |||   :::   ||||||            |||      :::   :::
Db        494 TrpGlyValLeuValPheSerThrSerAla---------GlyIleLeuTyrGlyThrIle 510

Qy       1460 CAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCC 1519
                 ::::||:::   |||   :::   ||||||      |||         |||
Db        511 PheThrValProPheIleLeuValAlaArgTyrHis---AlaLysAsnCysPheSerIle 529

Qy       1520 AAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTG 1579
              |||
Db        530 Lys--------------------------------------------------------- 530

Qy       1580 CCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGC 1639
                                           ||||||
Db        531 ---------------------------AsnGly--------------------------- 532

Qy       1640 CTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTG 1699

Db        532 ------------------------------------------------------------ 532
```

```
Alignment Scores:
Pred. No.:              4.73e-155     Length:         501
Score:                  2531.00       Matches:        489
Percent Similarity:     98.20%        Conservative:   2
Best Local Similarity:  97.80%        Mismatches:     9
Query Match:            39.44%        Indels:         0
DB:                     6             Gaps:           0

US-09-759-143-110 (1-3410) x Q95KC5 (1-501)

Qy        440 ATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGC 499
              |||||||||||||||||||||||||||||||||||||||||    |||||||||||||||
Db          1 MetThrMetValLeuGlyIleGlyProValLeuGlyLeuValSerValProLeuLeuGly 20

Qy        500 TCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTG 559
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 SerAlaSerAspHisTrpArgGlyArgTyrGlyArgArgArgProPheIleTrpAlaLeu 40

Qy        560 TCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTG 619
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 SerLeuGlyIleLeuLeuSerLeuPheLeuIleProArgAlaGlyTrpLeuAlaGlyLeu 60

Qy        620 CTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTG 679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LeuCysProAspProArgProLeuGluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeu 80

Qy        680 GACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGG 739
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 AspPheCysGlyGlnValCysPheThrProLeuGluAlaLeuLeuSerAspLeuPheArg 100

Qy        740 GACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGC 799
              ||||||||||||||||||||||||||||||||||||   |||||||||||||||||||||
Db        101 AspProAspHisCysArgGlnAlaTyrSerValTyrThrPheMetIleSerLeuGlyGly 120

Qy        800 TGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTG 859
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CysLeuGlyTyrLeuLeuProAlaIleAspTrpAspThrSerAlaLeuAlaProTyrLeu 140

Qy        860 GGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCA 919
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 GlyThrGlnGluGluCysLeuPheGlyLeuLeuThrLeuIlePheLeuThrCysValAla 160

Qy        920 GCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTG 979
              |||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db        161 AlaThrLeuLeuValAlaGluGluAlaAlaLeuGlyProAlaGluProAlaGluGlyLeu 180

Qy        980 TCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAAC 1039
              |||||||||||||||      |||||||||||||||   |||||||||||||||||||||
Db        181 SerAlaProSerLeuProSerHisCysCysProCysTrpAlaArgLeuAlaPheArgAsn 200

Qy       1040 CTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGC 1099
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 LeuGlyAlaLeuLeuProArgLeuHisGlnLeuCysCysArgMetProArgThrLeuArg 220

Qy       1100 CGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTAC 1159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 ArgLeuPheValAlaGluLeuCysSerTrpMetAlaLeuMetThrPheThrLeuPheTyr 240

Qy       1160 ACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAG 1219
              ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db        241 ThrAspPheValGlyGluGlyLeuTyrGlnGlyValProArgAlaGluLeuGlyThrGlu 260

Qy       1220 GCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGC 1279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 AlaArgArgHisTyrAspGluGlyValArgMetGlySerLeuGlyLeuPheLeuGlnCys 280

Qy       1280 GCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGA 1339
              |||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db        281 AlaIleSerLeuValPheSerLeuValValAspArgLeuValGlnArgPheGlyThrArg 300

Qy       1340 GCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCC 1399
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AlaValTyrLeuAlaSerValAlaAlaPheProValAlaAlaGlyAlaThrCysLeuSer 320

Qy       1400 CACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTG 1459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 HisSerValAlaValValThrAlaSerAlaAlaLeuThrGlyPheThrPheSerAlaLeu 340

Qy       1460 CAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCC 1519
              |||||||||||||||||||||||||||||||||||||||||||||||:::||||||||||
Db        341 GlnIleLeuProTyrThrLeuAlaSerLeuTyrHisArgGluArgGlnValPheLeuPro 360

Qy       1520 AAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTG 1579
              ||||||||||||||   ||||||   ||||||||||||||||||||||||||||||||||
Db        361 LysTyrArgGlyAspAlaGlyGlyThrSerSerGluAspSerLeuMetThrSerPheLeu 380

Qy       1580 CCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGC 1639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 ProGlyProLysProGlyAlaProPheProAsnGlyHisValGlyAlaGlyGlySerGly 400

Qy       1640 CTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTG 1699
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 LeuLeuProProProProAlaLeuCysGlyAlaSerAlaCysAspValSerValArgVal 420

Qy       1700 GTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTC 1759
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ValValGlyGluProThrGluAlaArgValValProGlyArgGlyIleCysLeuAspLeu 440

Qy       1760 GCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCC 1819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 AlaIleLeuAspSerAlaPheLeuLeuSerGlnValAlaProSerLeuPheMetGlySer 460

Qy       1820 ATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTG 1879
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 IleValGlnLeuSerGlnSerValThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeu 480

Qy       1880 GTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCA 1939
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ValAlaIleTyrPheAlaThrGlnValValPheAspLysSerAspLeuAlaLysTyrSer 500

RESULT 4
Q9VSV1
ID   Q9VSV1      PRELIMINARY;      PRT;   599 AA.
AC   Q9VSV1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   CG4484 protein.
GN   CG4484.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
```

```
Alignment Scores:
Pred. No.:              2.98e-172     Length:        553
Score:                  2798.00       Matches:       542
Percent Similarity:     98.55%        Conservative:  2
Best Local Similarity:  98.19%        Mismatches:    8
Query Match:            43.60%        Indels:        0
DB:                     6             Gaps:          0

US-09-759-143-110 (1-3410) x Q95KI5 (1-553)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 IleAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              |||||||||||||||||     ||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValSerValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||  ||||||||||||||||||||||||||||||||||      |||
Db        221 AlaAlaLeuGlyProAlaGluProAlaGluGlyLeuSerAlaProSerLeuProSerHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysTrpAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluLeuGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||:::||||||||||||||||||||||||||||||||    ||||||
Db        401 SerLeuTyrHisArgGluArgGlnValPheLeuProLysTyrArgGlyAspAlaGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ThrSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCA 1939
              ||||||||||||||||||||||||||||||||||||
Db        541 ValValPheAspLysSerAspLeuAlaLysTyrSer 552

RESULT 3
Q95KC5
ID   Q95KC5      PRELIMINARY;      PRT;    501 AA.
AC   Q95KC5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Hypothetical 53.4 kDa protein.
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=MEDULLA OBLOGATA;
RA   Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA   Suzuki Y., Sugano S., Hashimoto K.;
RT   "Isolation of full-length cDNA clones from macaque brain cDNA
RT   libraries.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB062977; BAB60745.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   501 AA;  53447 MW;  8C554BBD04EE0470 CRC64;
```

```
Pred. No.:                2.59e-176     Length:        553
Score:                    2861.00       Matches:       553
Percent Similarity:       100.00%       Conservative:  0
Best Local Similarity:    100.00%       Mismatches:    0
Query Match:              44.58%        Indels:        0
DB:                       4             Gaps:          0

US-09-759-143-110 (1-3410) x Q96JT2 (1-553)

Qy        284 ATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTG 343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetValGlnArgLeuTrpValSerArgLeuLeuArgHisArgLysAlaGlnLeuLeuLeu 20

Qy        344 GTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCG 403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 ValAsnLeuLeuThrPheGlyLeuGluValCysLeuAlaAlaGlyIleThrTyrValPro 40

Qy        404 CCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGT 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProLeuLeuLeuGluValGlyValGluGluLysPheMetThrMetValLeuGlyIleGly 60

Qy        464 CCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ProValLeuGlyLeuValCysValProLeuLeuGlySerAlaSerAspHisTrpArgGly 80

Qy        524 CGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTC 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ArgTyrGlyArgArgArgProPheIleTrpAlaLeuSerLeuGlyIleLeuLeuSerLeu 100

Qy        584 TTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 PheLeuIleProArgAlaGlyTrpLeuAlaGlyLeuLeuCysProAspProArgProLeu 120

Qy        644 GAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTC 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GluLeuAlaLeuLeuIleLeuGlyValGlyLeuLeuAspPheCysGlyGlnValCysPhe 140

Qy        704 ACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ThrProLeuGluAlaLeuLeuSerAspLeuPheArgAspProAspHisCysArgGlnAla 160

Qy        764 TACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCC 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TyrSerValTyrAlaPheMetIleSerLeuGlyGlyCysLeuGlyTyrLeuLeuProAla 180

Qy        824 ATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IleAspTrpAspThrSerAlaLeuAlaProTyrLeuGlyThrGlnGluGluCysLeuPhe 200

Qy        884 GGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAG 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GlyLeuLeuThrLeuIlePheLeuThrCysValAlaAlaThrLeuLeuValAlaGluGlu 220

Qy        944 GCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCAC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AlaAlaLeuGlyProThrGluProAlaGluGlyLeuSerAlaProSerLeuSerProHis 240

Qy       1004 TGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTG 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CysCysProCysArgAlaArgLeuAlaPheArgAsnLeuGlyAlaLeuLeuProArgLeu 260

Qy       1064 CACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 HisGlnLeuCysCysArgMetProArgThrLeuArgArgLeuPheValAlaGluLeuCys 280

Qy       1124 AGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTG 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerTrpMetAlaLeuMetThrPheThrLeuPheTyrThrAspPheValGlyGluGlyLeu 300

Qy       1184 TACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGC 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TyrGlnGlyValProArgAlaGluProGlyThrGluAlaArgArgHisTyrAspGluGly 320

Qy       1244 GTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTG 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 ValArgMetGlySerLeuGlyLeuPheLeuGlnCysAlaIleSerLeuValPheSerLeu 340

Qy       1304 GTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCA 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 ValMetAspArgLeuValGlnArgPheGlyThrArgAlaValTyrLeuAlaSerValAla 360

Qy       1364 GCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCT 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AlaPheProValAlaAlaGlyAlaThrCysLeuSerHisSerValAlaValValThrAla 380

Qy       1424 TCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCC 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 SerAlaAlaLeuThrGlyPheThrPheSerAlaLeuGlnIleLeuProTyrThrLeuAla 400

Qy       1484 TCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGT 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerLeuTyrHisArgGluLysGlnValPheLeuProLysTyrArgGlyAspThrGlyGly 420

Qy       1544 GCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AlaSerSerGluAspSerLeuMetThrSerPheLeuProGlyProLysProGlyAlaPro 440

Qy       1604 TTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTC 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 PheProAsnGlyHisValGlyAlaGlyGlySerGlyLeuLeuProProProProAlaLeu 460

Qy       1664 TGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCC 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CysGlyAlaSerAlaCysAspValSerValArgValValValGlyGluProThrGluAla 480

Qy       1724 AGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ArgValValProGlyArgGlyIleCysLeuAspLeuAlaIleLeuAspSerAlaPheLeu 500

Qy       1784 CTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTC 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 LeuSerGlnValAlaProSerLeuPheMetGlySerIleValGlnLeuSerGlnSerVal 520

Qy       1844 ACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAG 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 ThrAlaTyrMetValSerAlaAlaGlyLeuGlyLeuValAlaIleTyrPheAlaThrGln 540

Qy       1904 GTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCG 1942
              |||||||||||||||||||||||||||||||||||||||
Db        541 ValValPheAspLysSerAspLeuAlaLysTyrSerAla 553

RESULT 2
Q95KI5
ID   Q95KI5       PRELIMINARY;      PRT;   553 AA.
AC   Q95KI5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Hypothetical 59.4 kDa protein.
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TEMPORAL LOBE RIGHT;
RA   Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA   Suzuki Y., Sugano S., Hashimoto K.;
RT   "Isolation of full-length cDNA clones from macaque brain cDNA
RT   libraries.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB060851; BAB46871.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   553 AA;  59392 MW;  0718F3A91FB3BF1E CRC64;
```

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2003  Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:         June  9, 2003, 21:38:25 ; Search time 121.5 Seconds
                                          (without alignments)
                                          11565.772 Million cell updates/sec

Title:          US-09-759-143-110
Perfect score:  6418
Sequence:       1 gggaaccagcctgcacgcgc..........aaaaaaataaaaaaaaaaaa 3410

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters:      1343160

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09759143/runat_05062003_111319_23055/app_query.fasta_1.3591
-DB=SPTREMBL_21 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=45 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09759143_@CGN_1_1_184_@runat_05062003_111319_23055 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      SPTREMBL_21:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
                10:  sp_plant:*
                11:  sp_rodent:*
                12:  sp_virus:*
                13:  sp_vertebrate:*
                14:  sp_unclassified:*
                15:  sp_rvirus:*
                16:  sp_bacteriap:*
                17:  sp_archeap:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                SUMMARIES
                       %
Result                Query
  No.    Score   Match Length DB  ID                  Description
--------------------------------------------------------------------------
    1     2861    44.6    553  4  Q96JT2              Q96jt2 homo sapien
    2     2798    43.6    553  6  Q95KI5              Q95ki5 macaca fasc
    3     2531    39.4    501  6  Q95KC5              Q95kc5 macaca fasc
    4    519.5     8.1    599  5  Q9VSV1              Q9vsv1 drosophila
    5      512     8.0    112 11  Q8R1I0              Q8r1i0 mus musculu
    6      456     7.1    576 13  Q90Z74              Q90z74 oryzias lat
c   7    382.5     6.1    774 12  O41971              O41971 murid herpe
    8    366.5     5.7    612 10  Q9SQK5              Q9sqk5 vitis vinif
c   9    365.5     5.9    706 12  O41972              O41972 murid herpe
c  10      364     5.8   1414  5  Q26634              Q26634 strongyloce
c  11      361     5.8    608 12  Q9Q5K9              Q9q5k9 herpesvirus
c  12      359     5.8    727 12  O41973              O41973 murid herpe
c  13      357     5.7   1453 11  Q63079              Q63079 rattus norv
c  14    354.5     5.7   1445 13  O93251              O93251 rana catesb
   15    348.5     5.4    501 10  Q9SQK6              Q9sqk6 vitis vinif
   16    347.5     5.4    515 10  O65803              O65803 daucus caro
   17    347.5     5.4    515 10  Q9FNR6              Q9fnr6 daucus caro
c  18      347     5.6    890 12  Q8UZE1              Q8uze1 cercopithic
   19    346.5     5.4    515 10  Q9SLN7              Q9sln7 daucus caro
   20    345.5     5.4   1691 11  Q9ESQ2              Q9esq2 mus musculu
   21    344.5     5.4   5146  6  Q8SPM4              Q8spm4 bos taurus
c  22    342.5     5.5   1461  4  O76045              O76045 homo sapien
   23    338.5     5.3    501 10  Q9SP63              Q9sp63 vitis vinif
   24      338     5.3    512 10  Q9S7Z5              Q9s7z5 apium grave
   25    337.5     5.3    523 10  O04077              O04077 vicia faba
c  26    335.5     5.4   1549 11  Q60444              Q60444 cricetulus
c  27    335.5     5.4   1838 11  O88207              O88207 mus musculu
c  28      335     5.4   1840 11  Q9JI03              Q9ji03 rattus norv
c  29    334.5     5.4   1840 11  Q60467              Q60467 cricetulus
   30      333     5.2    512 10  Q9ZTB9              Q9ztb9 apium grave
   31      333     5.2    516 10  Q43653              Q43653 solanum tub
   32      333     5.2    524 10  Q9XHL6              Q9xhl6 pisum sativ
c  33      333     5.3   1378  5  O97405              O97405 haliotis di
c  34      333     5.3   1669 11  Q9QZS0              Q9qzs0 mus musculu
c  35      333     5.3   1838  4  Q15094              Q15094 homo sapien
   36    332.5     5.2    508 10  Q8VYX3              Q8vyx3 brassica ol
c  37    332.5     5.3   1487  6  O77753              O77753 canis famil
   38      332     5.2    502 10  Q9SP14              Q9sp14 alonsoa mer
c  39    330.5     5.3   1458 13  Q910B9              Q910b9 oncorhynchu
   40    328.5     5.1    506 10  Q944W2              Q944w2 oryza sativ
   41    328.5     5.1    594 10  O80605              O80605 arabidopsis
c  42      328     5.3   1418  6  Q28396              Q28396 equus cabal
   43    327.5     5.1    506 10  Q948L0              Q948l0 oryza sativ
   44    327.5     5.1    530 10  Q9M535              Q9m535 euphorbia e
   45    327.5     5.1    604 10  Q9FVL6              Q9fvl6 lycopersico

                                ALIGNMENTS

RESULT 1
Q96JT2
ID   Q96JT2      PRELIMINARY;      PRT;   553 AA.
AC   Q96JT2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Prostein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PROSTATE;
RX   MEDLINE=21139094; PubMed=11245466;
RA   Xu J., Kalos M., Stolk J.A., Zasloff E.J., Zhang X., Houghton R.L.,
RA   Filho A.M., Nolasco M., Badaro R., Reed S.G.;
RT   "Identification and characterization of prostein, a novel prostate-
RT   specific protein.";
RL   Cancer Res. 61:1563-1568(2001).
DR   EMBL; AY033593; AAK54386.1; -.
SQ   SEQUENCE   553 AA;  59322 MW;  0AFA23FBC742A667 CRC64;

Alignment Scores:
```

```
Qy       1417 GACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACAC 1476
              |||||||||||||||||||||||||||||| || ||||||||||||||||||||||||||
Db        215 GACAGCTTCAGCCGCCCTCACCGGGTTCACGTTTTCAGCCCTGCAGATCCTGCCCTACAC 156

Qy       1477 ACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCT 1515
              |||||||||||| ||||||||||||||||||||  || |
Db        155 ACTGGCCTCCCTATACCACCGGGAGAAGCAGGTAGTCAT 117

Search completed: June  9, 2003, 18:34:21
Job time : 8556 secs
```

```
RESULT 43
AR112290
LOCUS       AR112290                 272 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 11 from patent US 6130043.
ACCESSION   AR112290
VERSION     AR112290.1  GI:14092190
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 272)
  AUTHORS   Billing-Medel,P.A., Cohen,M., Colpitts,T.L., Friedman,P.N.,
            Gordon,J., Granados,E.N., Hodges,S.C., Klass,M.R., Kratochvil,J.D.,
            Roberts-Rapp,L., Russell,J.C., Stroupe,S.D. and Yu,H.
  TITLE     Reagents and methods useful for detecting diseases of the prostate
  JOURNAL   Patent: US 6130043-A 11 10-OCT-2000;
FEATURES             Location/Qualifiers
     source          1..272
                     /organism="unknown"
BASE COUNT       58 a     96 c     49 g     68 t      1 others
ORIGIN

  Query Match             7.9%;  Score 270;  DB 6;  Length 272;
  Best Local Similarity   99.6%;  Pred. No. 2.8e-36;
  Matches  270;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2795 AAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAA 2854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAA 60

Qy       2855 GGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTG 2914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTG 120

Qy       2915 GCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGA 2974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGA 180

Qy       2975 AGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCA 3034
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        181 AGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTNCCCTACCCCCAACTTTCCCCACCA 240

Qy       3035 GCTCCACAACCCTGTTTGGAGCTACTGCAGG 3065
              |||||||||||||||||||||||||||||||
Db        241 GCTCCACAACCCTGTTTGGAGCTACTGCAGG 271

RESULT 44
AR112288
LOCUS       AR112288                 265 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 9 from patent US 6130043.
ACCESSION   AR112288
VERSION     AR112288.1  GI:14092188
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 265)
  AUTHORS   Billing-Medel,P.A., Cohen,M., Colpitts,T.L., Friedman,P.N.,
            Gordon,J., Granados,E.N., Hodges,S.C., Klass,M.R., Kratochvil,J.D.,
            Roberts-Rapp,L., Russell,J.C., Stroupe,S.D. and Yu,H.
  TITLE     Reagents and methods useful for detecting diseases of the prostate
  JOURNAL   Patent: US 6130043-A 9 10-OCT-2000;
FEATURES             Location/Qualifiers
     source          1..265
                     /organism="unknown"
BASE COUNT       64 a     64 c     61 g     76 t
ORIGIN

  Query Match             7.8%;  Score 265;  DB 6;  Length 265;
  Best Local Similarity   100.0%;  Pred. No. 2e-35;
  Matches  265;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2453 GGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTC 2512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTC 60

Qy       2513 TTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGT 2572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGT 120

Qy       2573 TGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGG 2632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGG 180

Qy       2633 GAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCC 2692
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCC 240

Qy       2693 AACAATCAGGTCCCCTGAGATAGCT 2717
              |||||||||||||||||||||||||
Db        241 AACAATCAGGTCCCCTGAGATAGCT 265

RESULT 45
HSPE54C06/c
LOCUS       HSPE54C06                406 bp    DNA     linear   STS 18-NOV-1998
DEFINITION  H.sapiens flow-sorted chromosome 1 HindIII fragment, SC1pE54C06,
            sequence tagged site.
ACCESSION   AL033941
VERSION     AL033941.1  GI:3893713
KEYWORDS    STS; Anonymous marker; single read.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 406)
  AUTHORS   Gregory,S., Kettleborough,R., Langford,C., Ross,M.T. and Hunt,S.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-NOV-1998) The Sanger Centre, Wellcome Trust Genome
            Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
            humquery@sanger.ac.uk
COMMENT     Vector: pBSIISK+
            Marker stSG33426FS (Primer A : TGTAGGGCAGGATCTGCAG; Primer B :
            CTCGAGCAGTCTATTTGGCC; amplimer size : 141 bp) was mapped to
            chromosome 1 using Radiation Hybrid
            panel Genebridge 4 (GB4).
FEATURES             Location/Qualifiers
     source          1..406
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="SC1pE54C06"
                     /sex="Female"
                     /tissue_type="EBV lymphoblastoid cell line"
                     /clone_lib="SC1pE"
                     /dev_stage="adult"
BASE COUNT       97 a    128 c    115 g     66 t
ORIGIN

  Query Match             7.7%;  Score 261.4;  DB 11;  Length 406;
  Best Local Similarity   96.1%;  Pred. No. 8.4e-35;
  Matches  268;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1237 TGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTT 1296
              || |||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        395 TGCAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCGTGGTCTT 336

Qy       1297 CTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAG 1356
              || | |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        335 CTATGTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAG 276

Qy       1357 TGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGT 1416
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        275 TGTGGCAGCATTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGT 216
```

```
RESULT 40
AR112287
LOCUS       AR112287                 342 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 8 from patent US 6130043.
ACCESSION   AR112287
VERSION     AR112287.1  GI:14092187
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 342)
  AUTHORS   Billing-Medel,P.A., Cohen,M., Colpitts,T.L., Friedman,P.N.,
            Gordon,J., Granados,E.N., Hodges,S.C., Klass,M.R., Kratochvil,J.D.,
            Roberts-Rapp,L., Russell,J.C., Stroupe,S.D. and Yu,H.
  TITLE     Reagents and methods useful for detecting diseases of the prostate
  JOURNAL   Patent: US 6130043-A 8 10-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .342
                     /organism="unknown"
BASE COUNT       84 a     88 c     81 g     88 t      1 others
ORIGIN

  Query Match             9.4%;  Score 319.8;  DB 6;  Length 342;
  Best Local Similarity   98.5%;  Pred. No. 8.6e-45;
  Matches  333;  Conservative    0;  Mismatches    3;  Indels    2;  Gaps    1;

Qy       2207 GCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAG 2266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAG 60

Qy       2267 GGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATA 2326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATA 120

Qy       2327 AACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAG 2386
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAG 180

Qy       2387 CTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG--ACT 2444
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||  | |
Db        181 CTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGAAAGT 240

Qy       2445 TATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACA 2504
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACA 300

Qy       2505 GCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTT 2542
              |||||||||||||||||||||| ||||||||||| |||
Db        301 GCACTGTCTTTTTGCTGATCCANCCCCCTCTTACTTTT 338

RESULT 41
AR112291
LOCUS       AR112291                 294 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 12 from patent US 6130043.
ACCESSION   AR112291
VERSION     AR112291.1  GI:14092191
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 294)
  AUTHORS   Billing-Medel,P.A., Cohen,M., Colpitts,T.L., Friedman,P.N.,
            Gordon,J., Granados,E.N., Hodges,S.C., Klass,M.R., Kratochvil,J.D.,
            Roberts-Rapp,L., Russell,J.C., Stroupe,S.D. and Yu,H.
  TITLE     Reagents and methods useful for detecting diseases of the prostate
  JOURNAL   Patent: US 6130043-A 12 10-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .294
                     /organism="unknown"
BASE COUNT       70 a     87 c     60 g     74 t      3 others
ORIGIN

  Query Match             8.5%;  Score 291;  DB 6;  Length 294;
  Best Local Similarity   99.0%;  Pred. No. 7.1e-40;
  Matches  291;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2974 AAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACC 3033
              |||||||||||||||||  |||||||||||||||||||||||||||||||||||||||||
Db          1 AAGGCACTGCCCAAAATNNCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACC 60

Qy       3034 AGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAA 3093
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAA 120

Qy       3094 GCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACT 3153
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACT 180

Qy       3154 CAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCC 3213
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGGAGCACCCCCTGCNTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCC 240

Qy       3214 GTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 3267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 294

RESULT 42
AR112289
LOCUS       AR112289                 288 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 10 from patent US 6130043.
ACCESSION   AR112289
VERSION     AR112289.1  GI:14092189
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 288)
  AUTHORS   Billing-Medel,P.A., Cohen,M., Colpitts,T.L., Friedman,P.N.,
            Gordon,J., Granados,E.N., Hodges,S.C., Klass,M.R., Kratochvil,J.D.,
            Roberts-Rapp,L., Russell,J.C., Stroupe,S.D. and Yu,H.
  TITLE     Reagents and methods useful for detecting diseases of the prostate
  JOURNAL   Patent: US 6130043-A 10 10-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .288
                     /organism="unknown"
BASE COUNT       70 a     66 c     61 g     90 t      1 others
ORIGIN

  Query Match             8.4%;  Score 287;  DB 6;  Length 288;
  Best Local Similarity   99.7%;  Pred. No. 3.4e-39;
  Matches  287;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2532 CTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAG 2591
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAG 60

Qy       2592 GCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTT 2651
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTT 120

Qy       2652 CTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAG 2711
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db        121 CTGTGTTGGTGTCTAATATTTGGGTANGGTGGGGGATCCCCAACAATCAGGTCCCCTGAG 180

Qy       2712 ATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCA 2771
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCA 240

Qy       2772 AAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAA 2819
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAA 288
```

```
*     7819     7918: gap of unknown length
*     7919     9266: contig of 1348 bp in length
*     9267     9366: gap of unknown length
*     9367    10614: contig of 1248 bp in length
*    10615    10714: gap of unknown length
*    10715    11914: contig of 1200 bp in length
*    11915    12014: gap of unknown length
*    12015    14075: contig of 2061 bp in length
*    14076    14175: gap of unknown length
*    14176    15601: contig of 1426 bp in length
*    15602    15701: gap of unknown length
*    15702    17277: contig of 1576 bp in length
*    17278    17377: gap of unknown length
*    17378    18657: contig of 1280 bp in length
*    18658    18757: gap of unknown length
*    18758    20392: contig of 1635 bp in length
*    20393    20492: gap of unknown length
*    20493    21641: contig of 1149 bp in length
*    21642    21741: gap of unknown length
*    21742    23412: contig of 1671 bp in length
*    23413    23512: gap of unknown length
*    23513    25087: contig of 1575 bp in length
*    25088    25187: gap of unknown length
*    25188    26859: contig of 1672 bp in length
*    26860    26959: gap of unknown length
*    26960    28698: contig of 1739 bp in length
*    28699    28798: gap of unknown length
*    28799    30716: contig of 1918 bp in length
*    30717    30816: gap of unknown length
*    30817    32509: contig of 1693 bp in length
*    32510    32609: gap of unknown length
*    32610    34220: contig of 1611 bp in length
*    34221    34320: gap of unknown length
*    34321    35560: contig of 1240 bp in length
*    35561    35660: gap of unknown length
*    35661    37768: contig of 2108 bp in length
*    37769    37868: gap of unknown length
*    37869    39562: contig of 1694 bp in length
*    39563    39662: gap of unknown length
*    39663    41676: contig of 2014 bp in length
*    41677    41776: gap of unknown length
*    41777    43658: contig of 1882 bp in length
*    43659    43758: gap of unknown length
*    43759    45624: contig of 1866 bp in length
*    45625    45724: gap of unknown length
*    45725    48100: contig of 2376 bp in length
*    48101    48200: gap of unknown length
*    48201    50489: contig of 2289 bp in length
*    50490    50589: gap of unknown length
*    50590    52329: contig of 1740 bp in length
*    52330    52429: gap of unknown length
*    52430    54544: contig of 2115 bp in length
*    54545    54644: gap of unknown length
*    54645    56752: contig of 2108 bp in length
*    56753    56852: gap of unknown length
*    56853    58639: contig of 1787 bp in length
*    58640    58739: gap of unknown length
*    58740    60931: contig of 2192 bp in length
*    60932    61031: gap of unknown length
*    61032    63481: contig of 2450 bp in length
*    63482    63581: gap of unknown length
*    63582    66000: contig of 2419 bp in length
*    66001    66100: gap of unknown length
*    66101    68760: contig of 2660 bp in length
*    68761    68860: gap of unknown length
*    68861    71024: contig of 2164 bp in length
*    71025    71124: gap of unknown length
*    71125    73169: contig of 2045 bp in length
*    73170    73269: gap of unknown length
*    73270    76108: contig of 2839 bp in length
*    76109    76208: gap of unknown length
*    76209    79687: contig of 3479 bp in length
*    79688    79787: gap of unknown length
*    79788    82504: contig of 2717 bp in length
*    82505    82604: gap of unknown length
*    82605    86873: contig of 4269 bp in length
*    86874    86973: gap of unknown length
*    86974    91613: contig of 4640 bp in length
*    91614    91713: gap of unknown length
*    91714    94992: contig of 3279 bp in length
*    94993    95092: gap of unknown length
*    95093    97134: contig of 2042 bp in length
*    97135    97234: gap of unknown length
*    97235   100423: contig of 3189 bp in length
*   100424   100523: gap of unknown length
*   100524   104682: contig of 4159 bp in length
*   104683   104782: gap of unknown length
*   104783   108313: contig of 3531 bp in length
*   108314   108413: gap of unknown length
*   108414   113868: contig of 5455 bp in length
*   113869   113968: gap of unknown length
*   113969   117789: contig of 3821 bp in length
*   117790   117889: gap of unknown length
*   117890   120979: contig of 3090 bp in length
*   120980   121079: gap of unknown length
*   121080   123743: contig of 2664 bp in length
*   123744   123843: gap of unknown length

  Query Match             10.6%;  Score 361.8;  DB 2;  Length 198037;
  Best Local Similarity   77.3%;  Pred. No. 1e-51;
  Matches  486;  Conservative    0;  Mismatches  122;  Indels   21;  Gaps    3;

Qy          1505 CAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTG 1564
                 |||||||||||||||||||||||||||||| ||||||||| ||||||||| |||||||
Db         59507 CAGGTGTTCCTGCCCAAATACCGAGGGGACGCTGGAGGTGGTAGCAGTGAAGACAGCCAA 59448

Qy          1565 ATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGT 1624
                 |  ||||||||| ||| |||||||||||| |||||||||||||| ||||||||||||||
Db         59447 ACAACCAGCTTCTTGCTAGGCCCTAAGCCAGGAGCTCCCTTCCCCAATGGACACGTGGGC 59388

Qy          1625 GCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGAT 1684
                  |||| |||||  || | ||    || ||||| || ||||| ||||||||||||||||||
Db         59387 CCTGGCGGCAGCAGCATCCTGGTGCCCCCACCTGCACTCTGTGGGGCCTCTGCCTGTGAT 59328

Qy          1685 GTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGC 1744
                 |||||| | || ||||| |||||||||||  | || |||| ||| ||| | || ||||||
Db         59327 GTCTCCATGCGAGTGGTAGTGGGTGAGCCACCTGAAGCCAAGGTTGTTACTGGACGGGGC 59268

Qy          1745 ATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCC 1804
                 || ||||||||||| ||||||||||| |||||||| ||||||||||||||||| || |||
Db         59267 ATTTGCCTGGACCTTGCCATCCTGGACAGTGCCTTTCTGCTGTCCCAGGTGGCTCCGTCC 59208

Qy          1805 CTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCC 1864
                 ||||| |||||||||||||||||||| ||||| |||||||||||||||||||| || ||
Db         59207 CTGTTCATGGGCTCCATTGTCCAGCTGAGCCACTCTGTCACTGCCTATATGGTATCAGCT 59148

Qy          1865 GCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGAC 1924
                 |||||| ||||||||||||||||||||||||||||||||||||| ||||||||||  |||
Db         59147 GCAGGCTTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTGTTTGACAAGAATGAC 59088

Qy          1925 TTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACT 1984
                 |||||||||||||||  ||||||  ||   ||  |||    |  ||||  ||||||| |
Db         59087 TTGGCCAAATACTCACTGTAGAATTCTGTAAG-GCATCAAAGAAGAGGATCTGCCTCCCC 59029

Qy          1985 GGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCT 2044
                 || ||                 | |||||||  ||||||| | ||||| | |   ||||
Db         59028 GGTTC-----------------TCAGCCCCAGAGGGCTGCAGAGCTGGTCTC---TTTCC 58989

Qy          2045 GTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTG 2104
                 | |              |||||||||  |||||| ||   |    |||||||  | ||||
Db         58988 GGTCTCTGTTGCCCTGAGTGGCTCTCCACTGCCATCCGAAGGCAGTGAGGTGTATGGCTG 58929

Qy          2105 CACAGCTGGGGGCTGGGGCGTCCCTCTCC 2133
                 ||||| | || |||   |  ||   || |
Db         58928 CACAGGTTGGAGCTTTTGGCTCTGGCTTC 58900
```

```
Qy     2799 TCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAA 2858
            |||  |||||| || ||| |     ||                       | ||
Db     1109 TCTCTTCATCCTAATTGAGAGCCATAA-----------------------GATGTTCCTG 1145

Qy     2859 GGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCC 2918
               |  |  |||| |   ||||||| || || | || || | ||  |  ||||||| |||
Db     1146 CTCAAGGATTGGGACCCAAGGTCTCCACAGCATACCAAATC-CCAATTCTCTCTTGACCC 1204

Qy     2919 AGCCTGGTTCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGC 2978
            ||  ||||||||       ||||| | | |||                      ||| |
Db     1205 AGTTTGGTTCCCTGATTGCCCACTTCTCCCTA-------------------TGGGAAAGG 1245

Qy     2979 ACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTC 3038
            ||||   | |  || ||||                     |||| | |  || |  ||||
Db     1246 ACTGTGAAGAGATTGCCCT----------------------TCCAAGTCTGTCCCCTGGCTC 1285

Qy     3039 CACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTT 3098
            ||  | ||| | |   |    | | | | |  ||||||||| || |||| |  | || |
Db     1286 CAGGA-CCTCTCTATGGGGCTTTCTGAATCCTAAGCACAAA-TGTGGTTCCTTAGGCTTC 1343

Qy     3099 TGTCCATCTCAGCCCCCAGAGTATATCTGT----GCTTGGGGAATCTCACACAGAAACTC 3154
            || ||||||||||| |  | | ||||||||    || |||||||||||||||  ||||||
Db     1344 TGCCCATCTCAGCCTCTGGGGCATATCTGTCACAGCCTGGGGAATCTCACAC--AAACTC 1401

Qy     3155 AGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCG 3214
            ||||| || ||| ||| |||| |  |      | |    ||  ||||| ||| ||||| |
Db     1402 AGGAGTACTCCCCGCCAGAGCCAGTGAGCTCATCACG--TCGAGGGGGATTTCAGTGCTG 1459

Qy     3215 TTTGCAATAATGT------CGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAG 3268
             ||||||||||||       |||||||||||||||| | || |||||||||||| ||| |
Db     1460 ATTGCAATAATGTTATCTATGTCTTATTTATTTAGCAGAGTAAATATTTTATACCGTATG 1519

Qy     3269 TGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAA 3328
            |||  |||||||||||||||||||||||||  |||||||| || ||| ||||||||||||
Db     1520 TGATAAATCAGAGTATAATGTTTATGGTGATGAAATTAAAAGC-TTCGTATATGTTTAAA 1578

Qy     3329 AAAAAAAAAAAAAAA 3343
            |||||||||||||||
Db     1579 AAAAAAAAAAAAAAA 1593


RESULT 39
AC126523/c
LOCUS       AC126523                198037 bp    DNA     linear   HTG 14-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-65F14, *** SEQUENCING IN PROGRESS
            ***, 65 unordered pieces.
ACCESSION   AC126523
VERSION     AC126523.1  GI:21700488
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 198037)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 198037)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 198037)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GKBQ
                Center clone name: CH230-65F14
            -------------- Summary Statistics
                Sequencing vector: Plasmid;
                Chemistry: Dye-terminator Big Dye: 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 149884 bases at least Q40
                Consensus quality: 159599 bases at least Q30
                Consensus quality: 166549 bases at least Q20
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 65 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1063: contig of 1063 bp in length
            *     1064     1163: gap of unknown length
            *     1164     2325: contig of 1162 bp in length
            *     2326     2425: gap of unknown length
            *     2426     3556: contig of 1131 bp in length
            *     3557     3656: gap of unknown length
            *     3657     4752: contig of 1096 bp in length
            *     4753     4852: gap of unknown length
            *     4853     5972: contig of 1120 bp in length
            *     5973     6072: gap of unknown length
            *     6073     7818: contig of 1746 bp in length
```

```
Db         15 TTTGGGTAGGGTGGG 1

RESULT 38
BC024519
LOCUS       BC024519               1593 bp    mRNA    linear   ROD 07-AUG-2002
DEFINITION  Mus musculus, clone IMAGE:4984191, mRNA, partial cds.
ACCESSION   BC024519
VERSION     BC024519.1  GI:19353990
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1593)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAR-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 59 Row: p Column: 15
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis.
FEATURES             Location/Qualifiers
     source          1..1593
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /map="FVB/N"
                     /clone="IMAGE:4984191"
                     /tissue_type="Colon, normal. 5 month old male mouse."
                     /clone_lib="NCI_CGAP_Co24"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             <1..340
                     /codon_start=2
                     /product="Unknown (protein for IMAGE:4984191)"
                     /protein_id="AAH24519.1"
                     /db_xref="GI:19353991"
                     /translation="PNGHVGSGSSGILAPPPALCGASACDVSMRVVVGEPPEARVVTG
                     RGICLDLAILDSAFLLSQVAPSLFMGSIVQLSHSVTAYMVSAAGLGLVAIYFATQVVF
                     DKNDLAKYSV"
BASE COUNT      358 a    401 c    398 g    436 t
ORIGIN

  Query Match             13.0%;  Score 442.6;  DB 10;  Length 1593;
  Best Local Similarity   65.1%;  Pred. No. 1e-65;
  Matches 1142;  Conservative    0;  Mismatches  434;  Indels  179;  Gaps   26;

Qy       1606 CCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTG 1665
Db          1 CCCCAATGGACACGTGGGCTCTGGCAGCAGCGGCATCCTGGCCCCTCCACCTGCACTCTG 60

Qy       1666 CGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAG 1725
Db         61 TGGGGCCTCTGCCTGCGATGTCTCCATGCGAGTGGTGGTGGGTGAGCCACCTGAGGCCAG 120

Qy       1726 GGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCT 1785
Db        121 GGTTGTTACGGGACGGGGCATTTGCCTGGACCTGGCCATTCTGGACAGTGCCTTTCTGCT 180

Qy       1786 GTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCAC 1845
Db        181 GTCCCAGGTGGCTCCGTCCCTGTTCATGGGCTCCATTGTCCAGCTGAGCCACTCTGTCAC 240

Qy       1846 TGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGT 1905
Db        241 TGCCTATATGGTATCAGCTGCAGGCTTGGGTCTGGTCGCCATTTACTTTGCTACACAGGT 300

Qy       1906 AGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGG 1965
Db        301 AGTGTTTGACAAGAACGACTTGGCCAAATACTCAGTGTAGAATTGTGTAAG--GCATCAA 358

Qy       1966 GTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCC 2025
Db        359 AGAGAGGGTCTGCCTCATGGGTTCTCAGC---------------CCTTAGAGGGCTGCA 402

Qy       2026 GGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTG 2085
Db        403 GAGCTGGCCTCTCCAGGTCTTTGTCGCCTAA-----GTGGCTCTCTGCTGCCACCCTAAG 457

Qy       2086 CTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGT 2145
Db        458 GCAGTGAGGTGTATTGTTGCACAGATAGGAGCCAGAGCTT---TCGGGGCTCTGGCTTCA 514

Qy       2146 CTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGA 2205
Db        515 GAGTCTGGCTGGCCTACTGGCAGCCTCTCGCATGGAAGTCATTCTGAAGTTCTGCAGGGA 574

Qy       2206 GGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCA 2265
Db        575 T-------------------TGAAATGTGGTCTGTTCAGGTGTCCTGTATCAAGGCTCC 614

Qy       2266 GGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAAT 2325
Db        615 AGGTTAACAGCTAACCTCCTGCTTAAGT--CACCAGAGGGAGGATTTGGGGGACCTGAAC 672

Qy       2326 AAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTA 2385
Db        673 AAACTTCC---------------CTCCAAGTCCCTTGAC---CAGATTCTGCCACTGTT 713

Qy       2386 GCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTT 2445
Db        714 GCTCTTGCATGGGAGTTTCTTGTAGGATGCTCTCCTCCACAGGGTTTG---GCAGGGTGG 770

Qy       2446 ATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAG 2505
Db        771 AAGTTGGAAGGAGGAACCCTAGCAGGCAAGATGAAA---------CCCCGTGATCTACAG 821

Qy       2506 CACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTC 2565
Db        822 CACTGCCTTTCTGCTGAGTCTCCCACCCCCCACTTGTTTCAGGAAATGGCTCAGCTTGTT 881

Qy       2566 CTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA---CTTATTTATTTAACA 2622
Db        882 GGC---------CACAGAGACACAGGCCTTTAAATATTTAATTTATTATTTATTTAACA 931

Qy       2623 AAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTG 2682
Db        932 AATGAGAAGAAAATTCATTCCCAGCCTTTTTATGTT-GTGCCTAGGAGTTGAGTAGGGT- 989

Qy       2683 GGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGA 2742
Db        990 -GGGATTCCCCAAAACTGGGTCCCCTAAAATAGCTGGTCACTGTGTTAGATATTGCCAGA 1048

Qy       2743 ATCTTCTTCTCCTG----GGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAAT 2798
Db       1049 CTCCTCTTCTGGGGTGATAGATATGGTCCCCCAAACTGCTCAGCCCAGTACCTCGGGCAA 1108
```

```
JOURNAL    Patent: WO 0151633-A 11 19-JUL-2001;
           CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .772
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      221 a    202 c    170 g    162 t     17 others
ORIGIN

  Query Match             17.7%;  Score 603;  DB 6;  Length 772;
  Best Local Similarity   91.8%;  Pred. No. 3.6e-93;
  Matches  675;  Conservative    0;  Mismatches   49;  Indels   11;  Gaps    4;

Qy       1961 TTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCT-----GTTAGCCCCA 2015
              | ||||||| |||||||||| | | ||   |       | |   |      | | ||||
Db        735 TNGGGGTGGNGGGCCTGCCTNAATTGGGTTCCANGTTNCCCNGNTTCCTGTTAACCCCCN 676

Qy       2016 TGGGGCTGCCGGGCTGGCCGCC---AGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCT- 2071
              | ||| ||||||| | ||| ||   ||||| ||||  ||||||||||| |||| | | |
Db        675 TNGGGGTGCCGGGGTTGCCCCCCCNAGTTTTTGTTTGTGCCAAAGTAAAGTGGTTTTTTT 616

Qy       2072 GCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCT 2131
              ||||||||||||||||| |||||||||||| |||||||||||||| |||||||||| | |
Db        615 GCTGCCACCCTGTGCTGTTGAGGTGCGTAGNTGCACAGCTGGGGGNTGGGGCGTCCTTTT 556

Qy       2132 CCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTG 2191
              | | | |||||||| | ||||||||||||||||||| |||||||||||||||||||||||
Db        555 CTTTTTTCCCCAGTTTNTAGGGCTGCCTGACTGGAGNCCTTCCAAGGGGGTTTCAGTCTG 496

Qy       2192 GACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGG 2251
              ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        495 GACTTATACAGGGAGGCCAGAAGGGTTCCATGCACTGGAATGCGGGGACTCTGCAGGTGG 436

Qy       2252 ATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTT 2311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 ATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTT 376

Qy       2312 TTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAG 2371
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        375 TTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTNTAAGCCCCTTAACCTGCAG 316

Qy       2372 CTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATT 2431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 CTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATT 256

Qy       2432 TGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGG 2489
              ||||||||||  | ||||||||||||||||||||||||||||||||||||||||||||||
Db        255 TGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGG 196

Qy       2490 TCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGG 2549
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 TCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGG 136

Qy       2550 ATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACT 2609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 ATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACT 76

Qy       2610 TATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATA 2669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 TATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATA 16

Qy       2670 TTTGGGTAGGGTGGG 2684
              |||||||||||||||
Db         15 TTTGGGTAGGGTGGG 1


RESULT 37
AX267037/c
LOCUS       AX267037                772 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 11 from Patent WO0173032.
ACCESSION   AX267037
VERSION     AX267037.1  GI:16515822
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
            Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A., Hepler,W.T.
            and Henderson,R.A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0173032-A 11 04-OCT-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .772
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      221 a    202 c    170 g    162 t     17 others
ORIGIN

  Query Match             17.7%;  Score 603;  DB 6;  Length 772;
  Best Local Similarity   91.8%;  Pred. No. 3.6e-93;
  Matches  675;  Conservative    0;  Mismatches   49;  Indels   11;  Gaps    4;

Qy       1961 TTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCT-----GTTAGCCCCA 2015
              | ||||||| |||||||||| | | ||   |       | |   |      | | ||||
Db        735 TNGGGGTGGNGGGCCTGCCTNAATTGGGTTCCANGTTNCCCNGNTTCCTGTTAACCCCCN 676

Qy       2016 TGGGGCTGCCGGGCTGGCCGCC---AGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCT- 2071
              | ||| ||||||| | ||| ||   ||||| ||||  ||||||||||| |||| | | |
Db        675 TNGGGGTGCCGGGGTTGCCCCCCCNAGTTTTTGTTTGTGCCAAAGTAAAGTGGTTTTTTT 616

Qy       2072 GCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCT 2131
              ||||||||||||||||| |||||||||||| |||||||||||||| |||||||||| | |
Db        615 GCTGCCACCCTGTGCTGTTGAGGTGCGTAGNTGCACAGCTGGGGGNTGGGGCGTCCTTTT 556

Qy       2132 CCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTG 2191
              | | | |||||||| | ||||||||||||||||||| |||||||||||||||||||||||
Db        555 CTTTTTTCCCCAGTTTNTAGGGCTGCCTGACTGGAGNCCTTCCAAGGGGGTTTCAGTCTG 496

Qy       2192 GACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGG 2251
              ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        495 GACTTATACAGGGAGGCCAGAAGGGTTCCATGCACTGGAATGCGGGGACTCTGCAGGTGG 436

Qy       2252 ATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTT 2311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 ATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTT 376

Qy       2312 TTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAG 2371
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        375 TTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTNTAAGCCCCTTAACCTGCAG 316

Qy       2372 CTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATT 2431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 CTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATT 256

Qy       2432 TGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGG 2489
              ||||||||||  | ||||||||||||||||||||||||||||||||||||||||||||||
Db        255 TGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGG 196

Qy       2490 TCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGG 2549
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 TCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGG 136

Qy       2550 ATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACT 2609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 ATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACT 76

Qy       2610 TATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATA 2669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 TATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATA 16

Qy       2670 TTTGGGTAGGGTGGG 2684
```

```
Qy     2072 GCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGCTGGGGCGTCCCTCT 2131
Db      615 GCTGCCACCCTGTGCTGTTGAGGTGCGTAGNTGCACAGCTGGGGNTGGGGCGTCCTTTT 556

Qy     2132 CCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTG 2191
Db      555 CTTTTTTCCCCAGTTTNTAGGGCTGCCTGACTGGAGNCCTTCCAAGGGGGTTTCAGTCTG 496

Qy     2192 GACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGG 2251
Db      495 GACTTATACAGGGAGGCCAGAAGGGTTCCATGCACTGGAATGCGGGGACTCTGCAGGTGG 436

Qy     2252 ATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTT 2311
Db      435 ATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTT 376

Qy     2312 TTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAG 2371
Db      375 TTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTNTAAGCCCCTTAACCTGCAG 316

Qy     2372 CTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATT 2431
Db      315 CTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATT 256

Qy     2432 TGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGG 2489
Db      255 TGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGG 196

Qy     2490 TCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGG 2549
Db      195 TCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGG 136

Qy     2550 ATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACT 2609
Db      135 ATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACT 76

Qy     2610 TATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATA 2669
Db       75 TATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATA 16

Qy     2670 TTTGGGTAGGGTGGG 2684
Db       15 TTTGGGTAGGGTGGG 1
```

```
RESULT 35
AX140521/c
LOCUS       AX140521                 772 bp    DNA     linear   PAT 31-MAY-2001
DEFINITION  Sequence 11 from Patent WO0134802.
ACCESSION   AX140521
VERSION     AX140521.1  GI:14280639
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 772)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Retter,M.W., Stolk,J.A., Day,C.H.,
            Skeiky,Y.A. and Wang,A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0134802-A 11 17-MAY-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .772
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      221 a    202 c    170 g    162 t     17 others
ORIGIN

  Query Match             17.7%;  Score 603;  DB 6;  Length 772;
  Best Local Similarity   91.8%;  Pred. No. 3.6e-93;
  Matches  675;  Conservative    0;  Mismatches   49;  Indels   11;  Gaps    4;
```

```
Qy     1961 TTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCT-----GTTAGCCCCA 2015
Db      735 TNGGGGTGGNGGGCCTGCCTNAATTGGGTTCCANGTTNCCCNGNTTCCTGTTAACCCCCN 676

Qy     2016 TGGGGCTGCCGGGCTGGCCGCC---AGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCT- 2071
Db      675 TNGGGGTGCCGGGGTTGCCCCCCCNAGTTTTTGTTTGTGCCAAAGTAAAGTGGTTTTTTT 616

Qy     2072 GCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGCTGGGGCGTCCCTCT 2131
Db      615 GCTGCCACCCTGTGCTGTTGAGGTGCGTAGNTGCACAGCTGGGGNTGGGGCGTCCTTTT 556

Qy     2132 CCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTG 2191
Db      555 CTTTTTTCCCCAGTTTNTAGGGCTGCCTGACTGGAGNCCTTCCAAGGGGGTTTCAGTCTG 496

Qy     2192 GACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGG 2251
Db      495 GACTTATACAGGGAGGCCAGAAGGGTTCCATGCACTGGAATGCGGGGACTCTGCAGGTGG 436

Qy     2252 ATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTT 2311
Db      435 ATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTT 376

Qy     2312 TTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAG 2371
Db      375 TTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTNTAAGCCCCTTAACCTGCAG 316

Qy     2372 CTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATT 2431
Db      315 CTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATT 256

Qy     2432 TGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGG 2489
Db      255 TGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGG 196

Qy     2490 TCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGG 2549
Db      195 TCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGG 136

Qy     2550 ATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACT 2609
Db      135 ATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACT 76

Qy     2610 TATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATA 2669
Db       75 TATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATA 16

Qy     2670 TTTGGGTAGGGTGGG 2684
Db       15 TTTGGGTAGGGTGGG 1
```

```
RESULT 36
AX200381/c
LOCUS       AX200381                 772 bp    DNA     linear   PAT 29-AUG-2001
DEFINITION  Sequence 11 from Patent WO0151633.
ACCESSION   AX200381
VERSION     AX200381.1  GI:15390177
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 772)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Fanger,G.R., Day,C.H., Retter,M.W.,
            Stolk,J.A., Skeiky,Y.A., Wang,A. and Meagher,M.J.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
```

```
         *    73651    81298: contig of 7648 bp in length
         *    81299 81398: gap of      100 bp
         *    81399    91238: contig of 9840 bp in length
         *    91239 91338: gap of      100 bp
         *    91339    127752: contig of 36414 bp in length
         *   127753 127852: gap of      100 bp
         *   127853    177486: contig of 49634 bp in length
         *   177487 177586: gap of      100 bp
         *   177587    178567: contig of 981 bp in length.
FEATURES             Location/Qualifiers
     source          1..178567
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="RP23-272H13"
                     /clone_lib="RPCI-23 Female Mouse BAC"
     misc_feature    1..58509
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    58610..59459
                     /note="assembly_fragment"
     misc_feature    59560..60623
                     /note="assembly_fragment"
     misc_feature    60724..62950
                     /note="assembly_fragment"
     misc_feature    63051..65309
                     /note="assembly_fragment"
     misc_feature    65410..67839
                     /note="assembly_fragment"
     misc_feature    67940..69956
                     /note="assembly_fragment"
     misc_feature    70057..73550
                     /note="assembly_fragment"
     misc_feature    73651..81298
                     /note="assembly_fragment"
     misc_feature    81399..91238
                     /note="assembly_fragment"
     misc_feature    91339..127752
                     /note="assembly_fragment"
     misc_feature    127853..177486
                     /note="assembly_fragment"
     misc_feature    177587..178567
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
BASE COUNT    47757 a  40927 c  41360 g  47318 t   1205 others
ORIGIN

  Query Match             18.3%;  Score 623.6;  DB 2;  Length 178567;
  Best Local Similarity   86.8%;  Pred. No. 1.7e-96;
  Matches  686;  Conservative    0;  Mismatches  104;  Indels    0;  Gaps    0;

Qy        455 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 514
Db     171183 GGCATTGGCCCAGTGCTAGGCCTGGTTTCTGTTCCACTCCTAGGCTCAGCCAGTGACCAG 171124

Qy        515 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 574
Db     171123 TGGCGTGGGCGCTATGGCCGCCGGAGACCCTTTATCTGGGCTTTGTCCCTGGGTGTCCTG 171064

Qy        575 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 634
Db     171063 CTAAGCCTCTTTCTCATCCCGAGGGCTGGCTGGCTGGCAGGACTGCTGTACCCAGACACC 171004

Qy        635 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 694
Db     171003 AGGCCCCTGGAGTTGGCCCTGCTGATCTTGGGAGTGGGGCTGCTGGACTTTTGTGGCCAG 170944

Qy        695 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 754
Db     170943 GTGTGCTTTACTCCATTGGAGGCCTTACTCTCCGACCTCTTCCGGGACCCAGACCACTGC 170884

Qy        755 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 814
```

```
Db     170883 CGCCAAGCCTTCTCTGTCTACGCCTTCATGATCAGCCTTGGGGGCTGCCTGGGCTACCTC 170824

Qy        815 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 874
Db     170823 TTACCTGCCATTGACTGGGACACCAGCGTTCTGGCCCCCTACCTGGGTACTCAGGAAGAA 170764

Qy        875 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 934
Db     170763 TGCCTCTTTGGCCTCCTCACCCTCATTTTCCTCATCTGCATGGCAGCCACTCTGTTTGTG 170704

Qy        935 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 994
Db     170703 ACGGAGGAGGCAGTACTGGGCCCACCCGAGCCGGCAGAAGGGTTGTTGGTCTCTGCCGTG 170644

Qy        995 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1054
Db     170643 TCGCGCCGATGCTGCCCATGCCACGTTGGCCTGGCTTTCCGGAATCTGGGTACCCTGTTT 170584

Qy       1055 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1114
Db     170583 CCCCGGCTGCAGCAGCTGTGCTGCCGCATGCCTCGCACCCTACGCCGACTCTTTGTGGCT 170524

Qy       1115 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 1174
Db     170523 GAGCTGTGCAGCTGGATGGCACTTATGACTTTCACACTGTTCTACACGGACTTCGTGGGA 170464

Qy       1175 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 1234
Db     170463 GAGGGGCTGTACCAGGGTGTACCCAGAGCCGAGCCAGGCACCGAGGCCCGGAGACACTAT 170404

Qy       1235 GATGAAGGCG 1244
Db     170403 GATGAAGGTG 170394


RESULT 34
AX106230/c
LOCUS       AX106230               772 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 11 from Patent WO0125272.
ACCESSION   AX106230
VERSION     AX106230.1  GI:13921916
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 772)
  AUTHORS   Xu,J., Skeiky,Y.A., Reed,S.G. and Cheever,M.A.
  TITLE     Compositions and methods for therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0125272-A 11 12-APR-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..772
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    1..772
                     /note="n = A,T,C or G"
BASE COUNT      221 a    202 c    170 g    162 t     17 others
ORIGIN

  Query Match             17.7%;  Score 603;  DB 6;  Length 772;
  Best Local Similarity   91.8%;  Pred. No. 3.6e-93;
  Matches  675;  Conservative    0;  Mismatches   49;  Indels   11;  Gaps    4;

Qy       1961 TTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCT-----GTTAGCCCCA 2015
Db        735 TNGGGGTGGNGGGCCTGCCTNAATTGGGTTCCANGTTNCCCNGNTTCCTGTTAACCCCCN 676

Qy       2016 TGGGGCTGCCGGGCTGGCCGCC---AGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCT- 2071
Db        675 TNGGGGTGCCGGGGTTGCCCCCCCNAGTTTTTGTTTGTGCCAAAGTAAAGTGGTTTTTTT 616
```

```
Qy         875 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 934
               |||||||| ||||| ||||||||||| || |||| ||| || |||||||| ||||| |||
Db      161997 TGCCTCTTCGGCCTCCTCACCCTCATTTTTCTCATCTGTGTGGCAGCCACTCTGCTTGTG 162056

Qy         935 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 994
               |||||||||||||  || |||||  ||||||||||||||||| ||| || | ||||| ||
Db      162057 GCTGAGGAGGCAGTCCTTGGCCCACCCGAGCCAGCAGAAGGGTTGTTGGTCTCCTCCGTG 162116

Qy         995 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1054
               || | ||  |||||  ||||||  ||    |||||||||||||| |||||  ||||| ||
Db      162117 TCACGCCGGTGCTGCTCATGCCATGCTGGCCTGGCTTTCCGGAATCTGGGTACCCTGTTT 162176

Qy        1055 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1114
               |||||||||||||||||||||||||| ||||| |||||||||||||||||||| ||||| 
Db      162177 CCCCGGCTGCACCAGCTGTGCTGCCGAATGCCTCGCACCCTGCGCCGGCTCTTTGTGGCC 162236

Qy        1115 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 1174
               ||||||||||||||||||||||| ||||| ||||| ||||| |||||||| |||||||| 
Db      162237 GAGCTGTGCAGCTGGATGGCACTTATGACTTTCACACTGTTCTACACGGACTTCGTGGGA 162296

Qy        1175 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 1234
               ||||||||||||||||| || |||||||| ||||| || |||||||||||||||||||||
Db      162297 GAGGGGCTGTACCAGGGTGTCCCCAGAGCAGAGCCAGGTACCGAGGCCCGGAGACACTAT 162356

Qy        1235 GATGAAGGCG 1244
               |||||||| |
Db      162357 GATGAAGGTG 162366

RESULT 33
AC107837/c
LOCUS       AC107837              178567 bp    DNA     linear   HTG 14-APR-2002
DEFINITION  Mus musculus clone RP23-272H13, WORKING DRAFT SEQUENCE, 13 ordered
            pieces.
ACCESSION   AC107837
VERSION     AC107837.2  GI:20148020
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 178567)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus, clone RP23-272H13
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 178567)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C.,
            Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
            McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
            Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
            Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
            Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
            Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 178567)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
            Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
            Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-APR-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Apr 14, 2002 this sequence version replaced gi:18308540.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L20597
                Center clone name: 272_H_13
            -------------- Summary Statistics
                Sequencing vector: Plasmid; n/a; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 175041 bases at least Q40
                Consensus quality: 176485 bases at least Q30
                Consensus quality: 177028 bases at least Q20
                Insert size: 182000; agarose-fp
                Insert size: 177367; sum-of-contigs
                Quality coverage: 7.7 in Q20 bases; agarose-fp
                Quality coverage: 7.9 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 13 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *        1    58509: contig of 58509 bp in length
            *    58510 58609: gap of       100 bp
            *    58610    59459: contig of 850 bp in length
            *    59460 59559: gap of       100 bp
            *    59560    60623: contig of 1064 bp in length
            *    60624 60723: gap of       100 bp
            *    60724    62950: contig of 2227 bp in length
            *    62951 63050: gap of       100 bp
            *    63051    65309: contig of 2259 bp in length
            *    65310 65409: gap of       100 bp
            *    65410    67839: contig of 2430 bp in length
            *    67840 67939: gap of       100 bp
            *    67940    69956: contig of 2017 bp in length
            *    69957 70056: gap of       100 bp
            *    70057    73550: contig of 3494 bp in length
            *    73551 73650: gap of       100 bp
```

```
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1      1063: contig of 1063 bp in length
*     1064      1163: gap of unknown length
*     1164      2325: contig of 1162 bp in length
*     2326      2425: gap of unknown length
*     2426      3556: contig of 1131 bp in length
*     3557      3656: gap of unknown length
*     3657      4752: contig of 1096 bp in length
*     4753      4852: gap of unknown length
*     4853      5972: contig of 1120 bp in length
*     5973      6072: gap of unknown length
*     6073      7818: contig of 1746 bp in length
*     7819      7918: gap of unknown length
*     7919      9266: contig of 1348 bp in length
*     9267      9366: gap of unknown length
*     9367     10614: contig of 1248 bp in length
*    10615     10714: gap of unknown length
*    10715     11914: contig of 1200 bp in length
*    11915     12014: gap of unknown length
*    12015     14075: contig of 2061 bp in length
*    14076     14175: gap of unknown length
*    14176     15601: contig of 1426 bp in length
*    15602     15701: gap of unknown length
*    15702     17277: contig of 1576 bp in length
*    17278     17377: gap of unknown length
*    17378     18657: contig of 1280 bp in length
*    18658     18757: gap of unknown length
*    18758     20392: contig of 1635 bp in length
*    20393     20492: gap of unknown length
*    20493     21641: contig of 1149 bp in length
*    21642     21741: gap of unknown length
*    21742     23412: contig of 1671 bp in length
*    23413     23512: gap of unknown length
*    23513     25087: contig of 1575 bp in length
*    25088     25187: gap of unknown length
*    25188     26859: contig of 1672 bp in length
*    26860     26959: gap of unknown length
*    26960     28698: contig of 1739 bp in length
*    28699     28798: gap of unknown length
*    28799     30716: contig of 1918 bp in length
*    30717     30816: gap of unknown length
*    30817     32509: contig of 1693 bp in length
*    32510     32609: gap of unknown length
*    32610     34220: contig of 1611 bp in length
*    34221     34320: gap of unknown length
*    34321     35560: contig of 1240 bp in length
*    35561     35660: gap of unknown length
*    35661     37768: contig of 2108 bp in length
*    37769     37868: gap of unknown length
*    37869     39562: contig of 1694 bp in length
*    39563     39662: gap of unknown length
*    39663     41676: contig of 2014 bp in length
*    41677     41776: gap of unknown length
*    41777     43658: contig of 1882 bp in length
*    43659     43758: gap of unknown length
*    43759     45624: contig of 1866 bp in length
*    45625     45724: gap of unknown length
*    45725     48100: contig of 2376 bp in length
*    48101     48200: gap of unknown length
*    48201     50489: contig of 2289 bp in length
*    50490     50589: gap of unknown length
*    50590     52329: contig of 1740 bp in length
*    52330     52429: gap of unknown length
*    52430     54544: contig of 2115 bp in length
*    54545     54644: gap of unknown length
*    54645     56752: contig of 2108 bp in length
*    56753     56852: gap of unknown length
*    56853     58639: contig of 1787 bp in length
*    58640     58739: gap of unknown length
*    58740     60931: contig of 2192 bp in length
*    60932     61031: gap of unknown length
*    61032     63481: contig of 2450 bp in length
*    63482     63581: gap of unknown length
*    63582     66000: contig of 2419 bp in length
*    66001     66100: gap of unknown length
*    66101     68760: contig of 2660 bp in length
*    68761     68860: gap of unknown length
*    68861     71024: contig of 2164 bp in length
*    71025     71124: gap of unknown length
*    71125     73169: contig of 2045 bp in length
*    73170     73269: gap of unknown length
*    73270     76108: contig of 2839 bp in length
*    76109     76208: gap of unknown length
*    76209     79687: contig of 3479 bp in length
*    79688     79787: gap of unknown length
*    79788     82504: contig of 2717 bp in length
*    82505     82604: gap of unknown length
*    82605     86873: contig of 4269 bp in length
*    86874     86973: gap of unknown length
*    86974     91613: contig of 4640 bp in length
*    91614     91713: gap of unknown length
*    91714     94992: contig of 3279 bp in length
*    94993     95092: gap of unknown length
*    95093     97134: contig of 2042 bp in length
*    97135     97234: gap of unknown length
*    97235    100423: contig of 3189 bp in length
*   100424    100523: gap of unknown length
*   100524    104682: contig of 4159 bp in length
*   104683    104782: gap of unknown length
*   104783    108313: contig of 3531 bp in length
*   108314    108413: gap of unknown length
*   108414    113868: contig of 5455 bp in length
*   113869    113968: gap of unknown length
*   113969    117789: contig of 3821 bp in length
*   117790    117889: gap of unknown length
*   117890    120979: contig of 3090 bp in length
*   120980    121079: gap of unknown length
*   121080    123743: contig of 2664 bp in length
*   123744    123843: gap of unknown length

  Query Match             18.6%;  Score 633.2;  DB 2;  Length 198037;
  Best Local Similarity   87.6%;  Pred. No. 4e-98;
  Matches  692;  Conservative    0;  Mismatches   98;  Indels    0;  Gaps    0;

Qy        455 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 514
              |||||||| ||||||||||||||||| | ||| || ||||||||||||||||||||||| 
Db     161577 GGCATTGGCCCAGTGCTGGGCCTGGTTTCTGTTCCACTCCTAGGCTCAGCCAGTGACCAG 161636

Qy        515 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 574
              |||||||| ||||||||||||||  | ||||| |||||||| |||||| ||||  |||||
Db     161637 TGGCGTGGGCGCTATGGCCGCCGGAGACCCTTTATCTGGGCTCTGTCCCTGGGTGTCCTG 161696

Qy        575 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 634
              || |||||||| |||||||| |||||||||||||| |||||||| |||||  |  || | 
Db     161697 CTAAGCCTCTTCCTCATCCCGAGGGCCGGCTGGCTGGCAGGGCTACTGTGTTCAGATACT 161756

Qy        635 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 694
              |||||||||||| |||| ||||||||| |||| ||||||||||||||||| || ||||||
Db     161757 AGGCCCCTGGAGTTGGCCCTGCTCATCTTGGGAGTGGGGCTGCTGGACTTTTGCGGCCAG 161816

Qy        695 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 754
              |||||||| ||||||||||||||| | ||||| ||||||||||||||||| |||||||| 
Db     161817 GTGTGCTTTACTCCACTGGAGGCCTTACTCTCCGACCTCTTCCGGGACCCAGACCACTGC 161876

Qy        755 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 814
              ||||| |||| |||||||||||||||||||||||| || |||||||||||||||||||||
Db     161877 CGCCAAGCCTTCTCTGTCTATGCCTTCATGATCAGCCTCGGGGGCTGCCTGGGCTACCTC 161936

Qy        815 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 874
               | ||||||||||||||||||||||| ||||||||||||||||| ||||| ||||| || 
Db     161937 TTACCTGCCATTGACTGGGACACCAGCGCCCTGGCCCCCTACCTAGGCACTCAGGAAGAA 161996
```

```
Qy       1341 CAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCC 1400
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGTCTATNTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCC 60

Qy       1401 ACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGC 1460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGC 120

Qy       1461 AGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCA 1520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCA 180

Qy       1521 AATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC 1580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC 240

Qy       1581 CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCC 1640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCC 300

Qy       1641 TGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGG 1700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGG 360

Qy       1701 TGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCG 1760
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        361 TGGTGGGTGAGCCCACCGANGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCG 420

Qy       1761 CCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCA 1820
              ||||||||||||||| ||||||||||||||| ||||||||||||||||||||||||||||
Db        421 CCATCCTGGATAGTG-CTTCCTGCTGTCCCANGTGGCCCCATCCCTGTTTATGGGCTCCA 479

Qy       1821 TTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGG 1880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 TTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGG 539

Qy       1881 TCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAG 1940
              || ||||||||||||||||||||||| |||||||||||| ||| ||||||||||||||||
Db        540 TC-CCATTTACTTTGCTACACAGGTANTATTTGACAAGAACGANTTGGCCAAATACTCAG 598

Qy       1941 CGTAGAAAACTTCCAGCA--CATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCC 1998
              |||  |||| ||||||||    | |||||||| ||||||||||||||||| ||| |||||
Db        599 CGTTAAAAAATTCCAGCAACATTGGGGGTGGAAGGCCTGCCTCACTGGGT-CCAACTCCC 657

Qy       1999 CGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGT 2058
              ||||||||||| |||||||||||||||||  ||||||||| ||||||||||||||||| |
Db        658 CGCTCCTGTTAACCCCATGGGGCTGCCGGCTTGGCCGCCAATTTCTGTTGCTGCCAAANT 717

Qy       2059 AATGTGGCTCTCTGCTGCCACCCTGTGCT-GCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2117
               |||||||||||||||||||||   |||| ||||| |||| ||           ||||| 
Db        718 NATGTGGCTCTCTGCTGCCACCTGTTGCTGGCTGAAGTGCNTACNGCNCANCTNGGGGGG 777

Qy       2118 TGGGGCGTCCC 2128
              | ||| || ||
Db        778 TNGGGNGTTCC 788

RESULT 32
AC126523
LOCUS       AC126523              198037 bp    DNA     linear   HTG 14-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-65F14, *** SEQUENCING IN PROGRESS
            ***, 65 unordered pieces.
ACCESSION   AC126523
VERSION     AC126523.1  GI:21700488
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 198037)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 198037)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 198037)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GKBQ
                Center clone name: CH230-65F14
            -------------- Summary Statistics
                Sequencing vector: Plasmid;
                Chemistry: Dye-terminator Big Dye: 100% of reads
                Assembly program: Phrap; version 0.990329
                Consensus quality: 149884 bases at least Q40
                Consensus quality: 159599 bases at least Q30
                Consensus quality: 166549 bases at least Q20
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 65 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
```

```
                  ||||||||||||||| ||||||||||||||| ||||||||||||||||||||||||||||
Db          421 CCATCCTGGATAGTG-CTTCCTGCTGTCCCANGTGGCCCCATCCCTGTTTATGGCTCCA 479

Qy         1821 TTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGG 1880
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          480 TTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGG 539

Qy         1881 TCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAG 1940
                || ||||||||||||||||||||||| |||||||||||| || |||||||||||||||||
Db          540 TC-CCATTTACTTTGCTACACAGGTANTATTTGACAAGAACGANTTGGCCAAATACTCAG 598

Qy         1941 CGTAGAAAACTTCCAGCA--CATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCC 1998
                |||  |||| ||||||||    | |||||||| ||||||||||||||||| ||| |||||
Db          599 CGTTAAAAAATTCCAGCAACATTGGGGGTGGAAGGCCTGCCTCACTGGGT-CCAACTCCC 657

Qy         1999 CGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGT 2058
                ||||||||||| |||||||||||||||||  ||||||||| ||||||||||||||||| |
Db          658 CGCTCCTGTTAACCCCATGGGGCTGCCGGCTTGGCCGCCAATTTCTGTTGCTGCCAAANT 717

Qy         2059 AATGTGGCTCTCTGCTGCCACCCTGTGCT-GCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2117
                 |||||||||||||||||||||   |||| ||||| |||| ||           ||||| 
Db          718 NATGTGGCTCTCTGCTGCCACCTGTTGCTGGCTGAAGTGCNTACNGCNCANCTNGGGGG 777

Qy         2118 TGGGGCGTCCC 2128
                | ||| || ||
Db          778 TNGGGNGTTCC 788
```

RESULT 30
AX200380
LOCUS   AX200380   789 bp   DNA   linear   PAT 29-AUG-2001
DEFINITION Sequence 10 from Patent WO0151633.
ACCESSION   AX200380
VERSION   AX200380.1 GI:15390176
KEYWORDS   .
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 789)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Fanger,G.R., Day,C.H., Retter,M.W.,
            Stolk,J.A., Skeiky,Y.A., Wang,A. and Meagher,M.J.
  TITLE   Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0151633-A 10 19-JUL-2001;
            CORIXA CORPORATION (US)
FEATURES   Location/Qualifiers
     source   1..789
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT   122 a   250 c   221 g   182 t   14 others
ORIGIN

  Query Match   19.7%; Score 673.4; DB 6; Length 789;
  Best Local Similarity   94.2%; Pred. No. 3.3e-105;
  Matches 745; Conservative   0; Mismatches   40; Indels   6; Gaps   5;

```
Qy         1341 CAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCC 1400
                |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db            1 CAGTCTATNTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCC 60

Qy         1401 ACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGC 1460
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           61 ACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGC 120

Qy         1461 AGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCA 1520
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          121 AGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCA 180

Qy         1521 AATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC 1580
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          181 AATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC 240

Qy         1581 CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCC 1640
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          241 CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCC 300

Qy         1641 TGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGG 1700
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          301 TGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGG 360

Qy         1701 TGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCG 1760
                ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db          361 TGGTGGGTGAGCCCACCGANGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCG 420

Qy         1761 CCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGCTCCA 1820
                ||||||||||||||| ||||||||||||||| ||||||||||||||||||||||||||||
Db          421 CCATCCTGGATAGTG-CTTCCTGCTGTCCCANGTGGCCCCATCCCTGTTTATGGCTCCA 479

Qy         1821 TTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGG 1880
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          480 TTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGG 539

Qy         1881 TCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAG 1940
                || ||||||||||||||||||||||| |||||||||||| || |||||||||||||||||
Db          540 TC-CCATTTACTTTGCTACACAGGTANTATTTGACAAGAACGANTTGGCCAAATACTCAG 598

Qy         1941 CGTAGAAAACTTCCAGCA--CATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCC 1998
                |||  |||| ||||||||    | |||||||| ||||||||||||||||| ||| |||||
Db          599 CGTTAAAAAATTCCAGCAACATTGGGGGTGGAAGGCCTGCCTCACTGGGT-CCAACTCCC 657

Qy         1999 CGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGT 2058
                ||||||||||| |||||||||||||||||  ||||||||| ||||||||||||||||| |
Db          658 CGCTCCTGTTAACCCCATGGGGCTGCCGGCTTGGCCGCCAATTTCTGTTGCTGCCAAANT 717

Qy         2059 AATGTGGCTCTCTGCTGCCACCCTGTGCT-GCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2117
                 |||||||||||||||||||||   |||| ||||| |||| ||           ||||| 
Db          718 NATGTGGCTCTCTGCTGCCACCTGTTGCTGGCTGAAGTGCNTACNGCNCANCTNGGGGG 777

Qy         2118 TGGGGCGTCCC 2128
                | ||| || ||
Db          778 TNGGGNGTTCC 788
```

RESULT 31
AX267036
LOCUS   AX267036   789 bp   DNA   linear   PAT 26-OCT-2001
DEFINITION Sequence 10 from Patent WO0173032.
ACCESSION   AX267036
VERSION   AX267036.1 GI:16515821
KEYWORDS   .
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
            Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A., Hepler,W.T.
            and Henderson,R.A.
  TITLE   Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0173032-A 10 04-OCT-2001;
            CORIXA CORPORATION (US)
FEATURES   Location/Qualifiers
     source   1..789
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT   122 a   250 c   221 g   182 t   14 others
ORIGIN

  Query Match   19.7%; Score 673.4; DB 6; Length 789;
  Best Local Similarity   94.2%; Pred. No. 3.3e-105;
  Matches 745; Conservative   0; Mismatches   40; Indels   6; Gaps   5;

```
RESULT 28
AX106229
LOCUS       AX106229                 789 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 10 from Patent WO0125272.
ACCESSION   AX106229
VERSION     AX106229.1  GI:13921915
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 789)
  AUTHORS   Xu,J., Skeiky,Y.A., Reed,S.G. and Cheever,M.A.
  TITLE     Compositions and methods for therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0125272-A 10 12-APR-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .789
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    1. .789
                     /note="n = A,T,C or G"
BASE COUNT      122 a    250 c    221 g    182 t     14 others
ORIGIN

  Query Match             19.7%;  Score 673.4;  DB 6;  Length 789;
  Best Local Similarity   94.2%;  Pred. No. 3.3e-105;
  Matches  745;  Conservative    0;  Mismatches   40;  Indels    6;  Gaps    5;

Qy       1341 CAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCC 1400
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGTCTATNTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCC 60

Qy       1401 ACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGC 1460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGC 120

Qy       1461 AGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCA 1520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCA 180

Qy       1521 AATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC 1580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC 240

Qy       1581 CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCC 1640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCC 300

Qy       1641 TGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGG 1700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGG 360

Qy       1701 TGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCG 1760
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        361 TGGTGGGTGAGCCCACCGANGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCG 420

Qy       1761 CCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCA 1820
              ||||||||||||||| ||||||||||||||| ||||||||||||||||||||||||||||
Db        421 CCATCCTGGATAGTG-CTTCCTGCTGTCCCANGTGGCCCCATCCCTGTTTATGGGCTCCA 479

Qy       1821 TTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGG 1880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 TTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGG 539

Qy       1881 TCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAG 1940
              || ||||||||||||||||||||||| |||||||||||| ||| ||||||||||||||||
Db        540 TC-CCATTTACTTTGCTACACAGGTANTATTTGACAAGAACGANTTGGCCAAATACTCAG 598

Qy       1941 CGTAGAAAACTTCCAGCA--CATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCC 1998
              |||  ||||  |||||||        |  |||||||| |||||||||||||:|| |||||
Db        599 CGTTAAAAAATTCCAGCAACATTGGGGGTGGAAGGCCTGCCTCACTGGGT-CCAACTCCC 657
```

```
Qy       1999 CGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGT 2058
              ||||||||||| |||||||||||||||||  ||||||||| ||||||||||||||||| |
Db        658 CGCTCCTGTTAACCCCATGGGGCTGCCGGCTTGGCCGCCAATTTCTGTTGCTGCCAAANT 717

Qy       2059 AATGTGGCTCTCTGCTGCCACCCTGTGCT-GCTGAGGTGCGTAGCTGCACAGCTGGGGGC 2117
               ||||||||||||||||||||    |||| ||||| |||| ||          |||||
Db        718 NATGTGGCTCTCTGCTGCCACCTGTTGCTGGCTGAAGTGCNTACNGCNCANCTNGGGGGG 777

Qy       2118 TGGGGCGTCCC 2128
              | ||| || ||
Db        778 TNGGGNGTTCC 788


RESULT 29
AX140520
LOCUS       AX140520                 789 bp    DNA     linear   PAT 31-MAY-2001
DEFINITION  Sequence 10 from Patent WO0134802.
ACCESSION   AX140520
VERSION     AX140520.1  GI:14280638
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 789)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Retter,M.W., Stolk,J.A., Day,C.H.,
            Skeiky,Y.A. and Wang,A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0134802-A 10 17-MAY-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .789
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      122 a    250 c    221 g    182 t     14 others
ORIGIN

  Query Match             19.7%;  Score 673.4;  DB 6;  Length 789;
  Best Local Similarity   94.2%;  Pred. No. 3.3e-105;
  Matches  745;  Conservative    0;  Mismatches   40;  Indels    6;  Gaps    5;

Qy       1341 CAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCC 1400
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGTCTATNTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCC 60

Qy       1401 ACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGC 1460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGC 120

Qy       1461 AGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCA 1520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCA 180

Qy       1521 AATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC 1580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC 240

Qy       1581 CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCC 1640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCC 300

Qy       1641 TGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGG 1700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGG 360

Qy       1701 TGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCG 1760
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        361 TGGTGGGTGAGCCCACCGANGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCG 420

Qy       1761 CCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCA 1820
```

```
Qy         379 GGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGTAGAGGAGAAGTT 438
               ||||| |  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         402 GGCCGAATTCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGTAGAGGAGAAGTT 461

Qy         439 CATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGG 498
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         462 CATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGG 521

Qy         499 CTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACT 558
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         522 CTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACT 581

Qy         559 GTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCT 618
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         582 GTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCT 641

Qy         619 GCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCT 678
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         642 GCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCT 701

Qy         679 GGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCG 738
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         702 GGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCG 761

Qy         739 GGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGG 798
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         762 GGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGG 821

Qy         799 CTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCT 858
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         822 CTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCT 881

Qy         859 GGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGC 918
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         882 GGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGC 941

Qy         919 AGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCT 978
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         942 AGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCT 1001

Qy         979 GTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAA 1038
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1002 GTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAA 1061

Qy        1039 CCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCG 1098
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1062 CCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCG 1121

Qy        1099 CCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTA 1158
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1122 CCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTA 1181

Qy        1159 CACGGATTTCGTGGGCGAGGG 1179
               ||||||||||||||||||| |
Db        1182 CACGGATTTCGTGGGCGAGTG 1202
```

```
RESULT 27
AX267877
LOCUS       AX267877                1203 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 851 from Patent WO0173032.
ACCESSION   AX267877
VERSION     AX267877.1  GI:16516515
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
            Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A., Hepler,W.T.
            and Henderson,R.A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0173032-A 851 04-OCT-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..1203
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      175 a    415 c    368 g    245 t
ORIGIN

  Query Match             23.3%;  Score 794.6;  DB 6;  Length 1203;
  Best Local Similarity   99.5%;  Pred. No. 6.5e-126;
  Matches  797;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy         379 GGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGTAGAGGAGAAGTT 438
               ||||| |  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         402 GGCCGAATTCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTGGGGTAGAGGAGAAGTT 461

Qy         439 CATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGG 498
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         462 CATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGG 521

Qy         499 CTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACT 558
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         522 CTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACT 581

Qy         559 GTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCT 618
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         582 GTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCT 641

Qy         619 GCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCT 678
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         642 GCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCT 701

Qy         679 GGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCG 738
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         702 GGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCG 761

Qy         739 GGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGG 798
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         762 GGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGG 821

Qy         799 CTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCT 858
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         822 CTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCT 881

Qy         859 GGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGC 918
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         882 GGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGC 941

Qy         919 AGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCT 978
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         942 AGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCT 1001

Qy         979 GTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAA 1038
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1002 GTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAA 1061

Qy        1039 CCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCG 1098
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1062 CCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCG 1121

Qy        1099 CCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTA 1158
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1122 CCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTA 1181

Qy        1159 CACGGATTTCGTGGGCGAGGG 1179
               ||||||||||||||||||| |
Db        1182 CACGGATTTCGTGGGCGAGTG 1202
```

```
                |||||| | | | ||||||||    || |||||||||||||||  |||||||||| |
Db       2364 CTCAGCCTCTGGGGCATATCTGTCACAGCCTGGGGAATCTCACAC--AAACTCAGGAGTA 2421

Qy       3162 CCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAA 3221
                | ||| ||| |||| |  |          | |  ||      ||||| ||| ||||| | ||||||
Db       2422 CTCCCCGCCAGAGCCAGTGAGCTCATCACGTCGA--GGGGGATTTCAGTGCTGATTGCAA 2479

Qy       3222 TAATGT------CGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAA 3275
              ||||||      ||||||||||||||||| |  | |||||||||||  |||  ||||  ||
Db       2480 TAATGTTATCTATGTCTTATTTATTTAGCAGAGTAAATATTTTATACCGTATGTGATAAA 2539

Qy       3276 TCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAAAAAAA 3335
              ||||||||||||||||||||||  ||||||||| || ||| ||||||||||     |||||
Db       2540 TCAGAGTATAATGTTTATGGTGATGAAATTAAAAGC-TTCGTATATGTTTACTTGAAAAA 2598

Qy       3336 AAAAAAAAAAAAA 3348
              |||||||||||||
Db       2599 AAAAAAAAAAAAA 2611


RESULT 25
AX343860
LOCUS       AX343860                1593 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 6 from Patent WO0200892.
ACCESSION   AX343860
VERSION     AX343860.1  GI:18491927
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   cabezon Silva,T.E. and Delisse,A.M.
  TITLE     Triple fusion proteins comprising ubiquitin fused between
            thioredoxin and a polypeptide of interest
  JOURNAL   Patent: WO 0200892-A 6 03-JAN-2002;
            SmithKline Beecham Biologics SA (BE)
FEATURES             Location/Qualifiers
     source          1. .1593
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Chimaeric (E. coli - human)"
BASE COUNT      308 a    472 c    461 g    352 t
ORIGIN

  Query Match             28.2%;  Score 961.2;  DB 6;  Length 1593;
  Best Local Similarity   99.7%;  Pred. No. 2.2e-154;
  Matches  963;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        283 TATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCT 342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 TATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCT 662

Qy        343 GGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCC 402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 GGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCC 722

Qy        403 GCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGG 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 GCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGG 782

Qy        463 TCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGG 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 TCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGG 842

Qy        523 ACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCT 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 ACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCT 902

Qy        583 CTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCT 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        903 CTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCT 962

Qy        643 GGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTT 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 GGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTT 1022

Qy        703 CACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGC 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 CACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGC 1082

Qy        763 CTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1083 CTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGC 1142

Qy        823 CATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1143 CATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTT 1202

Qy        883 TGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGA 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1203 TGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGA 1262

Qy        943 GGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1263 GGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCA 1322

Qy       1003 CTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCT 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1323 CTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCT 1382

Qy       1063 GCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTG 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1383 GCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTG 1442

Qy       1123 CAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1443 CAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCT 1502

Qy       1183 GTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1503 GTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGG 1562

Qy       1243 CGTTCG 1248
              |  | |
Db       1563 CACTAG 1568


RESULT 26
AX201078
LOCUS       AX201078                1203 bp    DNA     linear   PAT 29-AUG-2001
DEFINITION  Sequence 708 from Patent WO0151633.
ACCESSION   AX201078
VERSION     AX201078.1  GI:15390883
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1203)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Fanger,G.R., Day,C.H., Retter,M.W.,
            Stolk,J.A., Skeiky,Y.A., Wang,A. and Meagher,M.J.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0151633-A 708 19-JUL-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .1203
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      175 a    415 c    368 g    245 t
ORIGIN

  Query Match             23.3%;  Score 794.6;  DB 6;  Length 1203;
  Best Local Similarity   99.5%;  Pred. No. 6.5e-126;
  Matches  797;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy    953 GGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCA 1012
Db    361 GGCCCACCCGAGCCGGCAGAAGGGTTGTTGGTCTCTGCCGTGTCGCGCCGATGCTGCCCA 420

Qy   1013 TGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTG 1072
Db    421 TGCCACGTTGGCCTGGCTTTCCGGAATCTGGGTACCCTGTTTCCCCGGCTGCAGCAGCTG 480

Qy   1073 TGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATG 1132
Db    481 TGCTGCCGCATGCCTCGCACCCTACGCCGACTCTTTGTGGCTGAGCTGTGCAGCTGGATG 540

Qy   1133 GCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGC 1192
Db    541 GCACTTATGACTTTCACACTGTTCTACACGGACTTCGTGGGAGAGGGGCTGTACCAGGGT 600

Qy   1193 GTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATG 1252
Db    601 GTACCCAGAGCCGAGCCAGGCACCGAGGCCCGGAGACACTATGATGAAGGCATTCGAATG 660

Qy   1253 GGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGAC 1312
Db    661 GGCAGCCTGGGGCTCTTCCTGCAGTGTGCCATCTCCCTGGTCTTCTCCCTGGTCATGGAC 720

Qy   1313 CGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCT 1372
Db    721 AGGCTGGTACAGAAGTTCGGCACACGGTCAGTCTATCTGGCCAGTGTGATGACCTTTCCT 780

Qy   1373 GTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCC 1432
Db    781 GTGGCTGCCGCTGCCACCTGCCTGTCCCACAGCGTGGTGGTAGTGACAGCCTCAGCTGCC 840

Qy   1433 CTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTAC 1492
Db    841 CTCACCGGGTTCACCTTCTCGGCCTTGCAGATCCTGCCTTACACGCTCGCCTCCCTCTAC 900

Qy   1493 CACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGT 1552
Db    901 CACCGTGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACGCTGGAGGTAGCAGCGGT 960

Qy   1553 GAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAAT 1612
Db    961 GAGGACAGCCAGACAACCAGCTTCTTGCCAGGCCCTAAGCCAGGAGCTCTCTTCCCCAAT 1020

Qy   1613 GGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCC 1672
Db   1021 GGACACGTGGGCTCTGGCAGCAGCGGCATCCTGGCCCCTCCACCTGCACTCTGTGGGGCC 1080

Qy   1673 TCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTT 1732
Db   1081 TCTGCCTGCGATGTCTCCATGCGAGTGGTGGTGGGTGAGCCACCTGAGGCCAGGGTTGTT 1140

Qy   1733 CCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAG 1792
Db   1141 ACGGGACGGGGCATTTGCCTGGACCTGGCCATTCTGGACAGTGCCTTTCTGCTGTCCCAG 1200

Qy   1793 GTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTAT 1852
Db   1201 GTGGCTCCGTCCCTGTTCATGGGCTCCATTGTCCAGCTGAGCCACTCTGTCACTGCCTAT 1260

Qy   1853 ATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTT 1912
Db   1261 ATGGTATCAGCTGCAGGCTTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTGTTT 1320

Qy   1913 GACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGG 1972
Db   1321 GACAAGAACGACTTGGCCAAATACTCAGTGTAGAATTGTGTAAGGCATCAAAGA--GAGG 1378

Qy   1973 GCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGG 2032
Db   1379 GTCTGCCTCATGGGTTCTCAGC----------------CCTTAGAGGGCTGCAGAGCTGG 1422

Qy   2033 CCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGA 2092
Db   1423 CCTCTCCAGGTCTTTGTCGCCTAA-----GTGGCTCTCTGCTGCCACCCTAAGGCAGTGA 1477

Qy   2093 GGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGG 2152
Db   1478 GGTGTATTGTTGCACAGATAGGAGCCAGAGCTT---TCGGGGCTCTGGCTTCAGAGTCTG 1534

Qy   2153 GCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGA 2212
Db   1535 GCTGGCCTACTGGCAGCCTCTCGCATGGAAGTCATTCTGAAGTTCTGCAGGG-------- 1586

Qy   2213 AGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAA 2272
Db   1587 ------------ATTGAAATGTGGTCTGTTCAGGTGTCCTGTATCAAGGCTCCAGGTTAA 1634

Qy   2273 CAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTGGGAGCTGAATAAACTCA 2332
Db   1635 CAGCTAACCTCCTGCTTAAG--TCACCAGAGGGAGGATTTGGGGGACCTGAACAAACT-- 1690

Qy   2333 GTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTG 2392
Db   1691 --------------TCCCTCCAAGTCCCTTGAC---CAGATTCTGCCACTGTTGCTCTTG 1733

Qy   2393 CATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGACTTATTTGTA 2452
Db   1734 CATGGGAGTTTCTTGTAGGATGCTCTCCTCCACAGGGTTTG---GCAGGGTGGAAGTTGG 1790

Qy   2453 GGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTC 2512
Db   1791 AAGGAGGAACCCTAGCAGGCAAGATGAAA---------CCCCGTGATCTACAGCACTGCC 1841

Qy   2513 TTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGT 2572
Db   1842 TTTCTGCTGAGTCTCCCACCCCCCACTTGTTTCAGGAAATGGCTCAGCTTGT-------- 1893

Qy   2573 TGCCATCACAGAGACACAGGCATTTAAATATTTAAC---TTATTTATTTAACAAAGTAGA 2629
Db   1894 --TGGCCACAGAGACACAGGCCTTTAAATATTTAATTTATTATTTATTTAACAAATGAGA 1951

Qy   2630 AGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATC 2689
Db   1952 AGAAAATTCATTCCCAGCCTTTTTATGTT-GTGCCTAGGAGTTGAGTAGGGT--GGGATT 2008

Qy   2690 CCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCT 2749
Db   2009 CCCCAAAACTGGGTCCCCTAAAATAGCTGGTCACTGTGTTAGATATTGCCAGACTCCTCT 2068

Qy   2750 TCTCCTGGGGTC----TGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTC 2805
Db   2069 TCTGGGGTGATAGATATGGTCCCCCAAACTGCTCAGCCCAGTACCTCGGGCAATCTCTTC 2128

Qy   2806 ATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAG 2865
Db   2129 ATCCTAATTGAGAGCCATAAGATGTTCCTGCTCAAG---------------------G 2165

Qy   2866 GGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGG 2925
Db   2166 ATTGGGACCCAAGGTCTCCACAGCATACCAAATCCCAATTC-TCTCTTGACCCAGTTTGG 2224

Qy   2926 TTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCC 2985
Db   2225 TTCCCTGATTGCCCACTTCTCCCTATGG------------------GAAAGGACTGTGA 2265

Qy   2986 AAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACC 3045
Db   2266 AGAGATTGCCCTT-------------------CCAAGTCTGTCCCCTGGCTCCAGGA-C 2304

Qy   3046 CTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCAT 3105
Db   2305 CTCTCTATGGGGCTTTCTGAATCCTAAGCACAAA-TGTGGTTCCTTAGGCTTCTGCCCAT 2363

Qy   3106 CTCAGCCCCCAGAGTATATCTGT----GCTTGGGGAATCTCACACAGAAACTCAGGAGCA 3161
```

```
Qy       1220 GCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGC 1279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1384 GCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGC 1443

Qy       1280 GCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGA 1339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1444 GCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGA 1503

Qy       1340 GCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCC 1399
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1504 GCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCC 1563

Qy       1400 CACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTG 1459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1564 CACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTG 1623

Qy       1460 CAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCC 1519
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1624 CAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCC 1683

Qy       1520 AAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTG 1579
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1684 AAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTG 1743

Qy       1580 CCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGC 1639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1744 CCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGC 1803

Qy       1640 CTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTG 1699
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1804 CTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTG 1863

Qy       1700 GTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTC 1759
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1864 GTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTC 1923

Qy       1760 GCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCC 1819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1924 GCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCC 1983

Qy       1820 ATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTG 1879
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1984 ATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTG 2043

Qy       1880 GTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCA 1939
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2044 GTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCA 2103

Qy       1940 GCG 1942
              |||
Db       2104 GCG 2106
```

```
RESULT 24
BC034084
LOCUS       BC034084                2611 bp    mRNA    linear   ROD 07-AUG-2002
DEFINITION  Mus musculus, clone IMAGE:4457452, mRNA, partial cds.
ACCESSION   BC034084
VERSION     BC034084.1  GI:21707615
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2611)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUL-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 43 Row: p Column: 14
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis.
FEATURES             Location/Qualifiers
     source          1..2611
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /map="FVB/N"
                     /clone="IMAGE:4457452"
                     /tissue_type="Mammary tumor. Metallothionien-TGF alpha
                     model. 10 month old virgin mouse. Taken by biopsy."
                     /clone_lib="NCI_CGAP_Mam1"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             <1..1353
                     /codon_start=1
                     /product="Unknown (protein for IMAGE:4457452)"
                     /protein_id="AAH34084.1"
                     /db_xref="GI:21707616"
                     /db_xref="LocusID:212980"
                     /translation="PRAGWLAGLLYPDTRPLELALLILGVGLLDFCGQVCFTPLEALL
                     SDLFRDPDHCRQAFSVYAFMISLGGCLGYLLPAIDWDTSVLAPYLGTQEECLFGLLTL
                     IFLICMAATLFVTEEAVLGPPEPAEGLLVSAVSRRCCPCHVGLAFRNLGTLFPRLQQL
                     CCRMPRTLRRLFVAELCSWMALMTFTLFYTDFVGEGLYQGVPRAEPGTEARRHYDEGI
                     RMGSLGLFLQCAISLVFSLVMDRLVQKFGTRSVYLASVMTFPVAAAATCLSHSVVVVT
                     ASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDAGGSSGEDSQTTSFLPGPKP
                     GALFPNGHVGSGSSGILAPPPALCGASACDVSMRVVVGEPPEARVVTGRGICLDLAIL
                     DSAFLLSQVAPSLFMGSIVQLSHSVTAYMVSAAGLGLVAIYFATQVVFDKNDLAKYSV
                     "
BASE COUNT      519 a    728 c    691 g    673 t
ORIGIN

  Query Match             36.5%;  Score 1246.2;  DB 10;  Length 2611;
  Best Local Similarity   73.1%;  Pred. No. 4.3e-203;
  Matches 2026;  Conservative    0;  Mismatches  568;  Indels  179;  Gaps   26;

Qy        593 CCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCA 652
              || ||||| |||||||| ||||| ||||||| ||| ||  |||||||||||||| ||||
Db          1 CCGAGGGCTGGCTGGCTGGCAGGACTGCTGTACCCAGACACCAGGCCCCTGGAGTTGGCC 60

Qy        653 CTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTG 712
              ||||| ||| |||| ||||||||||||||||| ||||||||||||||||| |||||| ||
Db         61 CTGCTGATCTTGGGAGTGGGGCTGCTGGACTTTTGTGGCCAGGTGTGCTTTACTCCATTG 120

Qy        713 GAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTC 772
              |||||| | ||||| ||||||||||||||||| |||||||| ||||| |||| |||||||
Db        121 GAGGCCTTACTCTCCGACCTCTTCCGGGACCCAGACCACTGCCGCCAAGCCTTCTCTGTC 180

Qy        773 TATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGG 832
              || |||||||||||||| |||||||||||||||||||||||| | |||||||||||||||
Db        181 TACGCCTTCATGATCAGCCTTGGGGGCTGCCTGGGCTACCTCTTACCTGCCATTGACTGG 240

Qy        833 GACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTC 892
              |||||||| |  ||||||||||||||||| || ||||| || |||||||||||||| |||
Db        241 GACACCAGCGTTCTGGCCCCCTACCTGGGTACTCAGGAAGAATGCCTCTTTGGCCTCCTC 300

Qy        893 ACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTG 952
              |||||||| ||||||| |||| | |||||||| ||| | ||| | ||||||||||  |||
Db        301 ACCCTCATTTTCCTCATCTGCATGGCAGCCACTCTGTTTGTGACGGAGGAGGCAGTACTG 360
```

```
Qy       2441 GACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCC 2500
Db       2533 GG---AAGTTGGAAGGAGGAACCCTAGCAGGCAAGATGAAA---------CCCCGTGATC 2580

Qy       2501 CACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGT 2560
Db       2581 TACAGCACTGCCTTTCTGCTGAGTCTCCCACCCCCCACTTGTTTCAGGAAATGGCTCAGC 2640

Qy       2561 TGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA---CTTATTTATT 2617
Db       2641 TTGTTGG---------CCACAGAGACACAGGCCTTTAAATATTTAATTTATTATTTATT 2690

Qy       2618 TAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTA 2677
Db       2691 TAACAAATGAGAAGAAAATTCATTCCCAGCCTTTTTATGTT-GTGCCTAGGAGTTGAGTA 2749

Qy       2678 GGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTG 2737
Db       2750 GGGT--GGGATTCCCCAAAACTGGGTCCCCTAAAATAGCTGGTCACTGTGTTAGATATTG 2807

Qy       2738 CCAGAATCTTCTTCTCCTG----GGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTG 2793
Db       2808 CCAGACTCCTCTTCTGGGGTGATAGATATGGTCCCCCAAACTGCTCAGCCCAGTACCTCG 2867

Qy       2794 GAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGA 2853
Db       2868 GGCAATCTCTTCATCCTAATTGAGAGCCATAA----------------------GATGT 2904

Qy       2854 AGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTT 2913
Db       2905 TCCTGCTCAAGGATTGGGACCCAAGGTCTCCACAGCATACCAAATC-CCAATTCTCTCTT 2963

Qy       2914 GGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATG 2973
Db       2964 GACCCAGTTTGGTTCCCTGATTGCCCACTTCTCCCTA------------------TGGG 3004

Qy       2974 AAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACC 3033
Db       3005 AAAGGACTGTGAAGAGATTGCCCT-------------------TCCAAGTCTGTCCCCT 3044

Qy       3034 AGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAA 3093
Db       3045 GGCTCCAGGA-CCTCTCTATGGGGCTTTCTGAATCCTAAGCACAAA-TGTGGTTCCTTAG 3102

Qy       3094 GCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGT----GCTTGGGGAATCTCACACAGA 3149
Db       3103 GCTTCTGCCCATCTCAGCCTCTGGGGCATATCTGTCACAGCCTGGGGAATCTCACAC--A 3160

Qy       3150 AACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAG 3209
Db       3161 AACTCAGGAGTACTCCCCGCCAGAGCCAGTGAGCTCATCAC--GTCGAGGGGGATTTCAG 3218

Qy       3210 TGCCGTTTGCAATAATGT------CGTCTTATTTATTTAGCGGGGTGAATATTTTATACT 3263
Db       3219 TGCTGATTGCAATAATGTTATCTATGTCTTATTTATTTAGCAGAGTAAATATTTTATACC 3278

Qy       3264 GTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGT 3323
Db       3279 GTATGTGATAAATCAGAGTATAATGTTTATGGTGATGAAATTAAAAGC-TTCGTATATGT 3337

Qy       3324 TTAAAAAAAAAAAAAA 3340
              ||||||||||||||||
Db       3338 TTAAAAAAAAAAAAAA 3354
```

```
RESULT 23
AX343857
LOCUS       AX343857                2133 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 3 from Patent WO0200892.
ACCESSION   AX343857
VERSION     AX343857.1  GI:18491926
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   cabezon Silva,T.E. and Delisse,A.M.
  TITLE     Triple fusion proteins comprising ubiquitin fused between
            thioredoxin and a polypeptide of interest
  JOURNAL   Patent: WO 0200892-A 3 03-JAN-2002;
            SmithKline Beecham Biologics SA (BE)
FEATURES             Location/Qualifiers
     source          1..2133
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Chimaeric (E. coli - human)"
BASE COUNT      384 a    656 c    614 g    479 t
ORIGIN

  Query Match             44.1%;  Score 1503;  DB 6;  Length 2133;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-247;
  Matches 1503;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        440 ATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGC 499
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 ATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGC 663

Qy        500 TCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTG 559
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 TCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTG 723

Qy        560 TCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTG 619
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        724 TCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTG 783

Qy        620 CTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTG 679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        784 CTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTG 843

Qy        680 GACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGG 739
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        844 GACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGG 903

Qy        740 GACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGC 799
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        904 GACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGC 963

Qy        800 TGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTG 859
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        964 TGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTG 1023

Qy        860 GGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCA 919
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1024 GGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCA 1083

Qy        920 GCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTG 979
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1084 GCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTG 1143

Qy        980 TCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAAC 1039
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1144 TCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAAC 1203

Qy       1040 CTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGC 1099
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1204 CTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGC 1263

Qy       1100 CGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTAC 1159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1264 CGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTAC 1323

Qy       1160 ACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAG 1219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1324 ACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAG 1383
```

```
Db    380 GCTGACTCAAAGCAGCTGTGGAGTATGTGAGTAGCCCCTGGAACCCTACCTGCCCTGTCC 439

Qy    281 ACTATGGTCCAGAGGCTGTGGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTG 340
Db    440 ATCATGATCCAGAGGCTGTGGGCCAGCCGTCTGCTACGGCACCGGAAAGCTCAGCTCCTG 499

Qy    341 CTGGTCAACCTGCTAACCTTTGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTG 400
Db    500 CTGGTCAACCTGCTCACCTTTGGCCTGGAGGTGTGCCTGGCTGCCGGCATTACCTATGTG 559

Qy    401 CCGCCTCTGCTGCTGGAAGTGGGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATT 460
Db    560 CCACCCCTTCTGCTGGAAGTCGGGGTGGAGGAGAAATTCATGACCATGGTGTTGGGCATT 619

Qy    461 GGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGT 520
Db    620 GGCCCAGTGCTAGGCCTGGTTTCTGTTCCACTCCTAGGCTCAGCCAGTGACCAGTGGCGT 679

Qy    521 GGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGC 580
Db    680 GGGCGCTATGGCCGCCGGAGACCCTTTATCTGGGCTTTGTCCCTGGGTGTCCTGCTAAGC 739

Qy    581 CTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCC 640
Db    740 CTCTTTCTCATCCCGAGGGCTGGCTGGCTGGCAGGACTGCTGTACCCAGACACCAGGCCC 799

Qy    641 CTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGC 700
Db    800 CTGGAGTTGGCCCTGCTGATCTTGGGAGTGGGGCTGCTGGACTTTTGTGGCCAGGTGTGC 859

Qy    701 TTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGTCGCCAG 760
Db    860 TTTACTCCATTGGAGGCCTTACTCTCCGACCTCTTCCGGGACCCAGACCACTGCCGCCAA 919

Qy    761 GCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCT 820
Db    920 GCCTTCTCTGTCTACGCCTTCATGATCAGCCTTGGGGGCTGCCTGGGCTACCTCTTACCT 979

Qy    821 GCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTC 880
Db    980 GCCATTGACTGGGACACCAGCGTTCTGGCCCCCTACCTGGGTACTCAGGAAGAATGCCTC 1039

Qy    881 TTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAG 940
Db   1040 TTTGGCCTCCTCACCCTCATTTTCCTCATCTGCATGGCAGCCACTCTGTTTGTGACGGAG 1099

Qy    941 GAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCC 1000
Db   1100 GAGGCAGTACTGGGCCCACCCGAGCCGGCAGAAGGGTTGTTGGTCTCTGCCGTGTCGCGC 1159

Qy   1001 CACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGG 1060
Db   1160 CGATGCTGCCCATGCCACGTTGGCCTGGCTTTCCGGAATCTGGGTACCCTGTTTCCCCGG 1219

Qy   1061 CTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTG 1120
Db   1220 CTGCAGCAGCTGTGCTGCCGCATGCCTCGCACCCTACGCCGACTCTTTGTGGCTGAGCTG 1279

Qy   1121 TGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGG 1180
Db   1280 TGCAGCTGGATGGCACTTATGACTTTCACACTGTTCTACACGGACTTCGTGGGAGAGGGG 1339

Qy   1181 CTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAA 1240
Db   1340 CTGTACCAGGGTGTACCCAGAGCCGAGCCAGGCACCGAGGCCCGGAGACACTATGATGAA 1399

Qy   1241 GGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCT 1300
Db   1400 GGCATTCGAATGGGCAGCCTGGGGCTCTTCCTGCAGTGTGCCATCTCCCTGGTCTTCTCC 1459

Qy   1301 CTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTG 1360
Db   1460 CTGGTCATGGACAGGCTGGTACAGAAGTTCGGCACACGGTCAGTCTATCTGGCCAGTGTG 1519

Qy   1361 GCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACA 1420
Db   1520 ATGACCTTTCCTGTGGCTGCCGCTGCCACCTGCCTGTCCCACAGCGTGGTGGTAGTGACA 1579

Qy   1421 GCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTG 1480
Db   1580 GCCTCAGCTGCCCTCACCGGGTTCACCTTCTCGGCCTTGCAGATCCTGCCTTACACGCTC 1639

Qy   1481 GCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGA 1540
Db   1640 GCCTCCCTCTACCACCGTGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACGCTGGA 1699

Qy   1541 GGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCT 1600
Db   1700 GGTAGCAGCGGTGAGGACAGCCAGACAACCAGCTTCTTGCCAGGCCCTAAGCCAGGAGCT 1759

Qy   1601 CCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCG 1660
Db   1760 CTCTTCCCCAATGGACACGTGGGCTCTGGCAGCAGCGGCATCCTGGCCCCTCCACCTGCA 1819

Qy   1661 CTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAG 1720
Db   1820 CTCTGTGGGGCCTCTGCCTGCGATGTCTCCATGCGAGTGGTGGTGGGTGAGCCACCTGAG 1879

Qy   1721 GCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTC 1780
Db   1880 GCCAGGGTTGTTACGGGACGGGGCATTTGCCTGGACCTGGCCATTCTGGACAGTGCCTTT 1939

Qy   1781 CTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCT 1840
Db   1940 CTGCTGTCCCAGGTGGCTCCGTCCCTGTTCATGGGCTCCATTGTCCAGCTGAGCCACTCT 1999

Qy   1841 GTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACA 1900
Db   2000 GTCACTGCCTATATGGTATCAGCTGCAGGCTTGGGTCTGGTCGCCATTTACTTTGCTACA 2059

Qy   1901 CAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACA 1960
Db   2060 CAGGTAGTGTTTGACAAGAACGACTTGGCCAAATACTCAGTGTAGAATTGTGTAAG--GC 2117

Qy   1961 TTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGG 2020
Db   2118 ATCAAAGAGAGGGTCTGCCTCATGGGTTCTCAGC----------------CCTTAGAGGG 2161

Qy   2021 CTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACC 2080
Db   2162 CTGCAGAGCTGGCCTCTCCAGGTCTTTGTCGCCTAA-----GTGGCTCTCTGCTGCCACC 2216

Qy   2081 CTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCC 2140
Db   2217 CTAAGGCAGTGAGGTGTATTGTTGCACAGATAGGAGCCAGAGCTT---TCGGGGCTCTGG 2273

Qy   2141 CCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATAC 2200
Db   2274 CTTCAGAGTCTGGCTGGCCTACTGGCAGCCTCTCGCATGGAAGTCATTCTGAAGTTCTGC 2333

Qy   2201 AGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAG 2260
Db   2334 AGGGAT--------------------TGAAATGTGGTCTGTTCAGGTGTCCTGTATCAAG 2373

Qy   2261 GCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTGGGAGC 2320
Db   2374 GCTCCAGGTTAACAGCTAACCTCCTGCTTAAGTCAC--CAGAGGGAGGATTTGGGGACC 2431

Qy   2321 TGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTA 2380
Db   2432 TGAACAAACTTCC----------------CTCCAAGTCCCTTGAC---CAGATTCTGCCA 2472

Qy   2381 ATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATAT 2440
Db   2473 CTGTTGCTCTTGCATGGGAGTTTCTTGTAGGATGCTCTCCTCCACAGGGTTTGGCAGGGT 2532
```

```
Qy          2430 TTTGAACATATGAC--TTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCA 2487
                 |||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        143150 TTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCA 143209

Qy          2488 GGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCA 2547
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143210 GGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCA 143269

Qy          2548 GGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA 2607
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143270 GGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA 143329

Qy          2608 CTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAA 2667
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143330 CTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAA 143389

Qy          2668 TATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGG 2727
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143390 TATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGG 143449

Qy          2728 CTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGG 2787
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143450 CTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGG 143509

Qy          2788 ACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGG 2847
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143510 ACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGG 143569

Qy          2848 TGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCT 2907
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143570 TGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCT 143629

Qy          2908 TCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGG 2967
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143630 TCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGG 143689

Qy          2968 CTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTC 3027
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143690 CTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTC 143749

Qy          3028 CCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 3087
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143750 CCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 143809

Qy          3088 TCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACA 3147
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143810 TCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACA 143869

Qy          3148 GAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTA 3207
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143870 GAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTA 143929

Qy          3208 AGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 3267
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143930 AGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 143989

Qy          3268 GTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTA 3326
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143990 GTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTA 144048
```

```
RESULT 22
BC031381
LOCUS       BC031381                3354 bp    mRNA    linear   ROD 07-AUG-2002
DEFINITION  Mus musculus, clone MGC:32471 IMAGE:5050610, mRNA, complete cds.
ACCESSION   BC031381
VERSION     BC031381.1  GI:21594808
KEYWORDS    MGC.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3354)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUN-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
            Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
            Richards, S., Gibbs, R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 45 Row: 1 Column: 2
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis.
FEATURES             Location/Qualifiers
     source          1. .3354
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /map="FVB/N"
                     /clone="MGC:32471 IMAGE:5050610"
                     /tissue_type="Liver, normal. 5 month old male mouse."
                     /clone_lib="NCI_CGAP_Li9"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             443. .2104
                     /codon_start=1
                     /product="Unknown (protein for MGC:32471)"
                     /protein_id="AAH31381.1"
                     /db_xref="GI:21594809"
                     /db_xref="LocusID:212980"
                     /translation="MIQRLWASRLLRHRKAQLLLVNLLTFGLEVCLAAGITYVPPLLL
                     EVGVEEKFMTMVLGIGPVLGLVSVPLLGSASDQWRGRYGRRRPFIWALSLGVLLSLFL
                     IPRAGWLAGLLYPDTRPLELALLILGVGLLDFCGQVCFTPLEALLSDLFRDPDHCRQA
                     FSVYAFMISLGGCLGYLLPAIDWDTSVLAPYLGTQEECLFGLLTLIFLICMAATLFVT
                     EEAVLGPPEPAEGLLVSAVSRRCCPCHVGLAFRNLGTLFPRLQQLCCRMPRTLRRLFV
                     AELCSWMALMTFTLFYTDFVGEGLYQGVPRAEPGTEARRHYDEGIRMGSLGLFLQCAI
                     SLVFSLVMDRLVQKFGTRSVYLASVMTFPVAAAATCLSHSVVVVTASAALTGFTFSAL
                     QILPYTLASLYHREKQVFLPKYRGDAGGSSGEDSQTTSFLPGPKPGALFPNGHVGSGS
                     SGILAPPPALCGASACDVSMRVVVGEPPEARVVTGRGICLDLAILDSAFLLSQVAPSL
                     FMGSIVQLSHSVTAYMVSAAGLGLVAIYFATQVVFDKNDLAKYSV"
BASE COUNT      662 a    922 c    934 g    836 t
ORIGIN

  Query Match             46.4%;  Score 1583.4;  DB 10;  Length 3354;
  Best Local Similarity   74.0%;  Pred. No. 1e-260;
  Matches 2456;  Conservative    0;  Mismatches  671;  Indels  190;  Gaps   28;

Qy            43 GCCTCGGCCAGGATCTGAGTGATGAGACGTGTCCCCACTGAG--GTGCCCCACAGCAGCA 100
                 | || | | |||||||||  || |||| |||||||||   ||    || | | |     |
Db           209 GGCTGGCCAAGGATCTGACCGACGAGATGTGTCCCCATCAAGCAAGGCACTAGATGGTGA 268

Qy           101 GGTGTTGAGCATGGGCTGAGAAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCC 160
                  ||||| ||| ||||  |||| |||| |  ||||| ||||||||||||||||||||  ||
Db           269 CGTGTTTAGCGTGGGACGAGATGCTGAATTGGCACTAAAGGGCTGGCAGAAATGGGAACC 328

Qy           161 TGGCTGATTCCTAGGCAGTTGGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGA 220
                 ||||||   ||||||  ||| | |  ||  ||||||||| ||| |         |||| 
Db           329 TGGCTGCACCCTAGGAGGTTAGTGCTAGTGAGGAGGAGAAGCCAC---------GGCAGG 379

Qy           221 GCCGAGACGAAGCAGTTCTGGAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCC 280
```

```
TGGCTTGCGGGCGTTCATGCTCAGAAACAAGGTTGGGACAAGCACTTCCAGGCTAACA
CAGTCAGAAATCGAAACGTACTCTCAACAGTTCGCTTAGGCATGGAAGTTTTGCGGCA
TTCTGGCTACACAATAACAAGGGAAGACTTACTCGTGGCTGCAACCCTACTAGCTCAA
AATTTATTCACACATGGTTACGCTTTGGGGAAATTATGAGGGGATCTCTCAGC
Phrap Quality:
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
BASE COUNT    40172 a  37516 c  36926 g  42774 t     600 others
ORIGIN

  Query Match             53.1%;  Score 1809.8;  DB 2;  Length 157988;
  Best Local Similarity   98.7%;  Pred. No. 2.8e-299;
  Matches 1835;  Conservative    0;  Mismatches   22;  Indels    2;  Gaps    1;

Qy       1470 CCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAG 1529
              |||   | ||  ||     ||| |  ||     |     ||||||||||||||||||||||
Db     142190 CCTGTCCTCTTCCCCTTTTCTTCACCCCTCTGCCTTAGGTGTTCCTGCCCAAATACCGAG 142249

Qy       1530 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTA 1589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142250 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTA 142309

Qy       1590 AGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCAC 1649
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142310 AGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCAC 142369

Qy       1650 CTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTG 1709
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142370 CTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTG 142429

Qy       1710 AGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGG 1769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142430 AGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGG 142489

Qy       1770 ATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGC 1829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142490 ATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGC 142549

Qy       1830 TCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTT 1889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142550 TCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTT 142609

Qy       1890 ACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAA 1949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142610 ACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAA 142669

Qy       1950 CTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTA 2009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142670 CTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTA 142729

Qy       2010 GCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCT 2069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142730 GCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCT 142789

Qy       2070 CTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT 2129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142790 CTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT 142849

Qy       2130 CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTC 2189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142850 CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTC 142909

Qy       2190 TGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT 2249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142910 TGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT 142969

Qy       2250 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG 2309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142970 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG 143029

Qy       2310 TTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGC 2369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143030 TTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGC 143089

Qy       2370 AGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGA 2429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143090 AGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGA 143149
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6432 ACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGG 6491

Qy       2848 TGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCT 2907
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6492 TGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCT 6551

Qy       2908 TCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGG 2967
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6552 TCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGG 6611

Qy       2968 CTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTC 3027
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6612 CTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTC 6671

Qy       3028 CCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 3087
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6672 CCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 6731

Qy       3088 TCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACA 3147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6732 TCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACA 6791

Qy       3148 GAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTA 3207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6792 GAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTA 6851

Qy       3208 AGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 3267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6852 AGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 6911

Qy       3268 GTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAA 3327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6912 GTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAA 6971

Qy       3328 AAAAA 3332
              |||||
Db       6972 AAAAA 6976
```

```
RESULT 21
AC096533
LOCUS       AC096533              157988 bp    DNA     linear   HTG 18-SEP-2001
DEFINITION  Homo sapiens chromosome 1 clone RP11-6B6, WORKING DRAFT SEQUENCE, 7
            unordered pieces.
ACCESSION   AC096533 AL365261
VERSION     AC096533.1  GI:15638681
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 157988)
  AUTHORS   Kaul,R.K., Olson,M.V., Raymond,C., Clendenning,J. and Haugen,E.D.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 157988)
  AUTHORS   Kaul,R.K., Olson,M.V., Raymond,C., Clendenning,J. and Haugen,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-SEP-2001) Genome Center, University of Washington,
            Box 352145, Seattle, WA 98195, USA
COMMENT     On Sep 18, 2001 this sequence version replaced gi:9931838.
            -------------- Genome Center
                Center: University of Washington Genome Center
                Center Code: UWGC
                Web site: http://www.genome.washington.edu
                Contact: uwgchtgs@u.washington.edu
                Drafting Center: SC
            -------------- Project Information
                Center project name: chr-1
                Center clone name: RP11-6B6 (sc0601)
            ------------- Summary Statistics
                Sequencing vector: plasmid; L08752; 100% of reads
                Chemistry: Dye-terminator ET; 57% of reads
                Chemistry: Dye-terminator Big Dye; 43% of reads
                Assembly program: Phrap; version 0.990319
                Consensus quality: 154688 bases at least Q40
                Consensus quality: 156347 bases at least Q30
                Consensus quality: 157058 bases at least Q20
                Insert size: 157388; sum-of-contigs
                Quality coverage: 7.3x in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 7 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *          1       6264: contig of 6264 bp in length
            *       6265       6364: gap of unknown length
            *       6365      15785: contig of 9421 bp in length
            *      15786      15885: gap of unknown length
            *      15886      27947: contig of 12062 bp in length
            *      27948      28047: gap of unknown length
            *      28048      51992: contig of 23945 bp in length
            *      51993      52092: gap of unknown length
            *      52093      81154: contig of 29062 bp in length
            *      81155      81254: gap of unknown length
            *      81255     118813: contig of 37559 bp in length
            *     118814     118913: gap of unknown length
            *     118914     157988: contig of 39075 bp in length.
FEATURES             Location/Qualifiers
     source          1..157988
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-6B6"
                     /clone_lib="RPCI human BAC library 11"
     misc_feature    1..6264
                     /note="assembly_name:Contig22"
     misc_feature    6365..15785
                     /note="assembly_name:Contig23"
     misc_feature    15886..27947
                     /note="assembly_name:Contig24"
     misc_feature    28048..51992
                     /note="assembly_name:Contig25"
     misc_feature    52093..81154
                     /note="assembly_name:Contig26"
     misc_feature    81255..118813
                     /note="assembly_name:Contig27"
     misc_feature    118914..157988
                     /note="assembly_name:Contig28"
     misc_feature    129829..129830
                     /note="gi|154843|gb|J01829|TRN10IS13 excised from sc0601
                     Contig28 ( 10917 . . 12245 )
                     TGATGAATCCCCTAATGATTTTGGTAAAAATCATTAAGTTAAGGTGGATACACATCTT
                     GTCATATGATCAAATGGTTTCGCGAAAAATCAATAATCAGACAACAAGATGTGCGAAC
                     TCGATATTTTACACGACTCTCTTTACCAATTCTGCCCCGAATTACACTTAAAACGACT
                     CAACAGCTTAACGTTGGCTTGCCACGCATTACTTGACTGTAAAACTCTCACTCTTACC
                     GAACTTGGCCGTAACCTGCCAACCAAAGCGAGAACAAAACATAACATCAAACGAATCG
                     ACCGATTGTTAGGTAATCGTCACCTCCACAAAGAGCGACTCGCTGTATACCGTTGGCA
                     TGCTAGCTTTATCTGTTCGGGCAATACGATGCCCATTGTACTTGTTGACTGGTCTGAT
                     ATTCGTGAGCAAAAACGACTTATGGTATTGCGAGCTTCAGTCGCACTACACGGTCGTT
                     CTGTTACTCTTTATGAGAAAGCGTTCCCGCTTTCAGAGCAATGTTCAAAGAAAGCTCA
                     TGACCAATTTCTAGCCGACCTTGCGAGCATTCTACCGAGTAACACCACACCGCTCATT
                     GTCAGTGATGCTGGCTTTAAAGTGCCATGGTATAAATCCGTTGAGAAGCTGGGTTGGT
                     ACTGGTTAAGTCGAGTAAGAGGAAAAGTACAATATGCAGACCTAGGAGCGGAAAACTG
                     GAAACCTATCAGCAACTTACATGATATGTCATCTAGTCACTCAAAGACTTTAGGCTAT
                     AAGAGGCTGACTAAAAGCAATCCAATCTCATGCCAAATTCTATTGTATAAATCTCGCT
                     CTAAAGGCCGAAAAAATCAGCGCTCGACACGGACTCATTGTCACCACCCGTCACCTAA
                     AATCTACTCAGCGTCGGCAAAGGAGCCATGGGTTCTAGCAACTAACTTACCTGTTGAA
                     ATTCGAACACCCAAACAACTTGTTAATATCTATTCGAAGCGAATGCAGATTGAAGAAA
                     CCTTCCGAGACTTGAAAAGTCCTGCCTACGGACTAGGCCTACGCCATAGCCGAACGAG
                     CAGCTCAGAGCGTTTTGATATCATGCTGCTAATCGCCCTGATGCTTCAACTAACATGT
```

```
Qy       3028 CCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 3087
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6672 CCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTT 6731

Qy       3088 TCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACA 3147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6732 TCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACA 6791

Qy       3148 GAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTA 3207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6792 GAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTA 6851

Qy       3208 AGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 3267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6852 AGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAA 6911

Qy       3268 GTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAA 3327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6912 GTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAA 6971

Qy       3328 AAAAA 3332
              |||||
Db       6972 AAAAA 6976

RESULT 20
AX267731
LOCUS       AX267731                6976 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 705 from Patent WO0173032.
ACCESSION   AX267731
VERSION     AX267731.1  GI:16516403
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
            Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A., Hepler,W.T.
            and Henderson,R.A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0173032-A 705 04-OCT-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..6976
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1312 a   1996 c   1961 g   1706 t      1 others
ORIGIN

  Query Match             53.2%;  Score 1815.8;  DB 6;  Length 6976;
  Best Local Similarity   98.7%;  Pred. No. 2e-300;
  Matches 1841;  Conservative    0;  Mismatches   22;  Indels    2;  Gaps    1;

Qy       1470 CCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAG 1529
              |||   | ||  ||    ||| |  ||    |    ||||||||||||||||||||||||
Db       5112 CCTGTCCTCTTCCCCTTTTCTTCACCCCTCTGCCTTAGGTGTTCCTGCCCAAATACCGAG 5171

Qy       1530 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTA 1589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5172 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTA 5231

Qy       1590 AGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCAC 1649
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5232 AGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCAC 5291

Qy       1650 CTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTG 1709
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5292 CTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTG 5351

Qy       1710 AGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGG 1769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5352 AGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGG 5411

Qy       1770 ATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGC 1829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5412 ATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGC 5471

Qy       1830 TCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTT 1889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5472 TCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTT 5531

Qy       1890 ACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAA 1949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5532 ACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAA 5591

Qy       1950 CTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTA 2009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5592 CTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTA 5651

Qy       2010 GCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCT 2069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5652 GCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCT 5711

Qy       2070 CTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT 2129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5712 CTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT 5771

Qy       2130 CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTC 2189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5772 CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTC 5831

Qy       2190 TGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT 2249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5832 TGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT 5891

Qy       2250 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG 2309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5892 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG 5951

Qy       2310 TTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGC 2369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5952 TTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGC 6011

Qy       2370 AGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGA 2429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6012 AGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGA 6071

Qy       2430 TTTGAACATATGAC--TTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCA 2487
              |||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db       6072 TTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCA 6131

Qy       2488 GGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCA 2547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6132 GGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCA 6191

Qy       2548 GGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA 2607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6192 GGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA 6251

Qy       2608 CTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAA 2667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6252 CTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAA 6311

Qy       2668 TATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGG 2727
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6312 TATTTGGGTAGGGTGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGG 6371

Qy       2728 CTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGG 2787
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6372 CTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGG 6431

Qy       2788 ACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGG 2847
```

```
Qy       3215 TTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCA 3274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2337 TTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCA 2396

Qy       3275 ATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAAAAAA 3334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2397 ATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAAAAAA 2456

Qy       3335 AAAAAAAAAAAAAAAAAAAAA 3355
              |||||||||||||||||||||
Db       2457 AAAAAAAAAAAAAAAAAAAAA 2477


RESULT 19
AX200996
LOCUS       AX200996                6976 bp    DNA     linear   PAT 29-AUG-2001
DEFINITION  Sequence 626 from Patent WO0151633.
ACCESSION   AX200996
VERSION     AX200996.1  GI:15390823
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6976)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Fanger,G.R., Day,C.H., Retter,M.W.,
            Stolk,J.A., Skeiky,Y.A., Wang,A. and Meagher,M.J.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0151633-A 626 19-JUL-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..6976
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1312 a   1996 c   1961 g   1706 t      1 others
ORIGIN

  Query Match             53.2%;  Score 1815.8;  DB 6;  Length 6976;
  Best Local Similarity   98.7%;  Pred. No. 2e-300;
  Matches 1841;  Conservative    0;  Mismatches   22;  Indels    2;  Gaps    1;

Qy       1470 CCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAG 1529
              |||   | ||  ||    ||| |  ||     |    ||||||||||||||||||||||||
Db       5112 CCTGTCCTCTTCCCCTTTTCTTCACCCCTCTGCCTTAGGTGTTCCTGCCCAAATACCGAG 5171

Qy       1530 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTA 1589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5172 GGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTA 5231

Qy       1590 AGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCAC 1649
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5232 AGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCAC 5291

Qy       1650 CTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTG 1709
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5292 CTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTG 5351

Qy       1710 AGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGG 1769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5352 AGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGG 5411

Qy       1770 ATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGC 1829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5412 ATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGC 5471

Qy       1830 TCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTT 1889
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5472 TCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTT 5531

Qy       1890 ACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAA 1949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5532 ACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAA 5591

Qy       1950 CTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTA 2009
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5592 CTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTA 5651

Qy       2010 GCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCT 2069
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5652 GCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCT 5711

Qy       2070 CTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT 2129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5712 CTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCT 5771

Qy       2130 CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTC 2189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5772 CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTC 5831

Qy       2190 TGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT 2249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5832 TGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGT 5891

Qy       2250 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG 2309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5892 GGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGG 5951

Qy       2310 TTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGC 2369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5952 TTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGC 6011

Qy       2370 AGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGA 2429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6012 AGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGA 6071

Qy       2430 TTTGAACATATGAC--TTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCA 2487
              |||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db       6072 TTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCA 6131

Qy       2488 GGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCA 2547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6132 GGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCA 6191

Qy       2548 GGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA 2607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6192 GGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAA 6251

Qy       2608 CTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAA 2667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6252 CTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAA 6311

Qy       2668 TATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGG 2727
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6312 TATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGG 6371

Qy       2728 CTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGG 2787
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6372 CTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGG 6431

Qy       2788 ACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGG 2847
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6432 ACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGG 6491

Qy       2848 TGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCT 2907
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6492 TGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCT 6551

Qy       2908 TCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGG 2967
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6552 TCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGG 6611

Qy       2968 CTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTC 3027
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6612 CTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTC 6671
```

```
                /clone="DKFZp666D0110"
                /tissue_type="stomach"
                /clone_lib="666 (synonym: hsto2). Vector pSport1; host
                DH10B; sites NotI + SalI"
                /dev_stage="adult"
     polyA_signal    2425..2430
     polyA_site      2447
BASE COUNT      505 a    713 c    634 g    625 t
ORIGIN

  Query Match             61.7%;  Score 2103.8;  DB 9;  Length 2477;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 2117;  Conservative    0;  Mismatches    2;  Indels    2;  Gaps    1;

Qy       1237 TGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTT 1296
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 TGAGGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTT 416

Qy       1297 CTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAG 1356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        417 CTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAG 476

Qy       1357 TGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGT 1416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        477 TGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGT 536

Qy       1417 GACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACAC 1476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 GACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACAC 596

Qy       1477 ACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACAC 1536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 ACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACAC 656

Qy       1537 TGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGG 1596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGG 716

Qy       1597 AGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACC 1656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 AGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACC 776

Qy       1657 CGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCAC 1716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 CGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCAC 836

Qy       1717 CGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGC 1776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        837 CGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGC 896

Qy       1777 CTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCA 1836
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        897 CTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCA 956

Qy       1837 GTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGC 1896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        957 GTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGC 1016

Qy       1897 TACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAG 1956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1017 TACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAG 1076

Qy       1957 CACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCAT 2016
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1077 CACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCAT 1136

Qy       2017 GGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGC 2076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1137 GGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGC 1196

Qy       2077 CACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCT 2136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1197 CACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCT 1256

Qy       2137 CTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTT 2196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1257 CTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTT 1316

Qy       2197 ATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTAC 2256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1317 ATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTAC 1376

Qy       2257 CCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGG 2316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1377 CCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGG 1436

Qy       2317 GAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCG 2376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1437 GAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCG 1496

Qy       2377 TTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAAC 2436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1497 TTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAAC 1556

Qy       2437 ATATG--ACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCC 2494
              |||||  | |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1557 ATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCC 1616

Qy       2495 TCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTG 2554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1617 TCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTG 1676

Qy       2555 GCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTT 2614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1677 GCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTT 1736

Qy       2615 ATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGG 2674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1737 ATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGG 1796

Qy       2675 GTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCA 2734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1797 GTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCA 1856

Qy       2735 TTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGG 2794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1857 TTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGG 1916

Qy       2795 AAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAA 2854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1917 AAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAA 1976

Qy       2855 GGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTG 2914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1977 GGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTG 2036

Qy       2915 GCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGA 2974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2037 GCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGA 2096

Qy       2975 AGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCA 3034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2097 AGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCA 2156

Qy       3035 GCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAG 3094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2157 GCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAG 2216

Qy       3095 CCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTC 3154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2217 CCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTC 2276

Qy       3155 AGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCG 3214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2277 AGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCG 2336
```

```
Qy       1605 TCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCT 1664
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCT 480

Qy       1665 GCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCA 1724
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCA 540

Qy       1725 GGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGC 1784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGC 600

Qy       1785 TGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCA 1844
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCA 660

Qy       1845 CTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGG 1904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGG 720

Qy       1905 TAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGG 1964
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGG 780

Qy       1965 GGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGC 2024
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGC 840

Qy       2025 CGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGT 2084
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGT 900

Qy       2085 GCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAG 2144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAG 960

Qy       2145 TCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGGTTTCAGTCTGGACTTATACAGGG 2204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGGTTTCAGTCTGGACTTATACAGGG 1020

Qy       2205 AGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTC 2264
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db       1021 AGGCCAGAAGGGCTCCATGCACTGGAATGC-GGGACTCTGCAGGTGGATTACCCAGGCTC 1079

Qy       2265 AGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAA 2324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 AGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAA 1139

Qy       2325 TAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGT 2384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 TAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGT 1199

Qy       2385 AGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG--A 2442
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |
Db       1200 AGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATGAAA 1259

Qy       2443 CTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCA 2502
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCA 1319

Qy       2503 CAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTG 2562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 CAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTG 1379

Qy       2563 GTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACA 2622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 GTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACA 1439

Qy       2623 AAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTG 2682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1440 AAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTG 1499

Qy       2683 GGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGA 2742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1500 GGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGA 1559

Qy       2743 ATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTA 2802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1560 ATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTA 1619

Qy       2803 CTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTA 2862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 CTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTA 1679

Qy       2863 GAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCC 2922
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1680 GAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCC 1739

Qy       2923 TGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTG 2982
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 TGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTG 1799

Qy       2983 CCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACA 3042
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 CCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACA 1859

Qy       3043 ACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTC 3102
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1860 ACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTC 1919

Qy       3103 CATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCAC 3162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1920 CATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCAC 1979

Qy       3163 CCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAAT 3222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1980 CCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAAT 2039

Qy       3223 AATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGT 3282
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       2040 AATGTCGTCTTATTTTTTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGT 2099

Qy       3283 ATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTA 3326
              ||||||||||||||||||||||||||||||||||||||||||||
Db       2100 ATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTA 2143
```

```
RESULT 18
HSM804244
LOCUS       HSM804244               2477 bp    mRNA    linear   PRI 10-JUL-2002
DEFINITION  Homo sapiens mRNA; cDNA DKFZp666D0110 (from clone DKFZp666D0110).
ACCESSION   AL832933
VERSION     AL832933.1  GI:21733520
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2477)
  AUTHORS   Ansorge,W., Wirkner,U., Mewes,H.W., Weil,B. and Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by EMBL (European Molecular Biology Laboratories,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            This clone (DKFZp666D0110) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://mips.gsf.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1..2477
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CACATTGGGGTGGAGGGCCTGCCTCACTGGTCCCAGCTCCCCGCTCCTGTTAGCCCCAT 840

Qy       2017 GGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGC 2076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGC 900

Qy       2077 CACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCT 2136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CACCCTGTGCTGCTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCT 960

Qy       2137 CTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTT 2196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTT 1020

Qy       2197 ATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTAC 2256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTAC 1080

Qy       2257 CCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGG 2316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGG 1140

Qy       2317 GAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCG 2376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCG 1200

Qy       2377 TTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAAC 2436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAAC 1260

Qy       2437 ATATG--ACTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCC 2494
              |||||  | |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ATATGAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCC 1320

Qy       2495 TCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTG 2554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTG 1380

Qy       2555 GCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTT 2614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTT 1440

Qy       2615 ATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGG 2674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ATTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGG 1500

Qy       2675 GTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCA 2734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GTAGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCA 1560

Qy       2735 TTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGG 2794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TTGCCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGG 1620

Qy       2795 AAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAA 2854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AAATTCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAA 1680

Qy       2855 GGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTG 2914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GGAAGGTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTG 1740

Qy       2915 GCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGA 2974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GCCCAGCCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGA 1800

Qy       2975 AGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCA 3034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 AGGCACTGCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCA 1860

Qy       3035 GCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAG 3094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GCTCCACAACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAG 1920

Qy       3095 CCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTC 3154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CCTTTGTCCATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTC 1980

Qy       3155 AGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCG 3214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 AGGAGCACCCCCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCG 2040

Qy       3215 TTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCA 3274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TTTGCAATAATGTCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCA 2100

Qy       3275 ATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTA 3326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 ATCAGAGTATAATGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTA 2152
```

```
RESULT 17
AR112294
LOCUS       AR112294                 2143 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 15 from patent US 6130043.
ACCESSION   AR112294
VERSION     AR112294.1  GI:14092194
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2143)
  AUTHORS   Billing-Medel,P.A., Cohen,M., Colpitts,T.L., Friedman,P.N.,
            Gordon,J., Granados,E.N., Hodges,S.C., Klass,M.R., Kratochvil,J.D.,
            Roberts-Rapp,L., Russell,J.C., Stroupe,S.D. and Yu,H.
  TITLE     Reagents and methods useful for detecting diseases of the prostate
  JOURNAL   Patent: US 6130043-A 15 10-OCT-2000;
FEATURES             Location/Qualifiers
     source          1..2143
                     /organism="unknown"
BASE COUNT      418 a    621 c    563 g    541 t
ORIGIN

  Query Match             62.0%;  Score 2114.8;  DB 6;  Length 2143;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 2139;  Conservative    0;  Mismatches    2;  Indels    3;  Gaps    2;

Qy       1185 ACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCG 1244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGCG 60

Qy       1245 TTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGG 1304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGG 120

Qy       1305 TCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAG 1364
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAG 180

Qy       1365 CTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTT 1424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTT 240

Qy       1425 CAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCT 1484
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCT 300

Qy       1485 CCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTG 1544
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTG 360

Qy       1545 CTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCT 1604
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCT 420
```

```
Db       3891 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCT 3950

Qy       2389 CTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTA 2446
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||  | |||
Db       3951 CTTGCATGGGAGTTTCTAGGATGAAACACTCCACCATGGGATTTGAACATATGAAAGTTA 4010

Qy       2447 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 2506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4011 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 4070

Qy       2507 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 2566
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4071 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 4130

Qy       2567 TTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGT 2626
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4131 TTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGT 4190

Qy       2627 AGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 2686
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4191 AGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 4250

Qy       2687 ATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCT 2746
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4251 ATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCT 4310

Qy       2747 TCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCA 2806
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4311 TCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCA 4370

Qy       2807 TCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGG 2866
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4371 TCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGG 4430

Qy       2867 GTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGT 2926
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4431 GTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGT 4490

Qy       2927 TCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 2986
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4491 TCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 4550

Qy       2987 AAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCC 3046
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4551 AAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCC 4610

Qy       3047 TGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATC 3106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4611 TGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATC 4670

Qy       3107 TCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCC 3166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4671 TCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCC 4730

Qy       3167 TGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATG 3226
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4731 TGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATG 4790

Qy       3227 TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAA 3286
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4791 TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAA 4850

Qy       3287 TGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAA 3330
              ||||||||||||||||||||||||||||||||||||||||||||
Db       4851 TGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAA 4894
```

```
RESULT 16
AR112295
LOCUS       AR112295                2152 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 16 from patent US 6130043.
ACCESSION   AR112295
VERSION     AR112295.1  GI:14092195
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2152)
  AUTHORS   Billing-Medel,P.A., Cohen,M., Colpitts,T.L., Friedman,P.N.,
            Gordon,J., Granados,E.N., Hodges,S.C., Klass,M.R., Kratochvil,J.D.,
            Roberts-Rapp,L., Russell,J.C., Stroupe,S.D. and Yu,H.
  TITLE     Reagents and methods useful for detecting diseases of the prostate
  JOURNAL   Patent: US 6130043-A 16 10-OCT-2000;
FEATURES             Location/Qualifiers
     source          1..2152
                     /organism="unknown"
BASE COUNT      419 a    622 c    569 g    542 t
ORIGIN

  Query Match             62.7%;  Score 2136.4;  DB 6;  Length 2152;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 2149;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    1;

Qy       1177 GGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGA 1236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGA 60

Qy       1237 TGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTT 1296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTT 120

Qy       1297 CTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAG 1356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAG 180

Qy       1357 TGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGT 1416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGT 240

Qy       1417 GACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACAC 1476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACAC 300

Qy       1477 ACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACAC 1536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACAC 360

Qy       1537 TGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGG 1596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGG 420

Qy       1597 AGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACC 1656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACC 480

Qy       1657 CGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCAC 1716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCAC 540

Qy       1717 CGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGC 1776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGC 600

Qy       1777 CTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCA 1836
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCA 660

Qy       1837 GTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGC 1896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGC 720

Qy       1897 TACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAG 1956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAG 780

Qy       1957 CACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCAT 2016
```

```
Qy    875 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 934
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1731 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 1790

Qy    935 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 994
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1791 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 1850

Qy    995 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1054
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1851 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1910

Qy   1055 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1114
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1911 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1970

Qy   1115 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 1174
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1971 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 2030

Qy   1175 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 1234
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2031 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 2090

Qy   1235 GATGA------------------------------------------------------- 1239
          |||||
Db   2091 GATGAAGGTAAGGCCTTGGCAGCCAGCAGAGGCTGGTGTGGGAGCCGCCCACCAGAGACG 2150

Qy   1240 ------------------------------------------------------------ 1239

Db   2151 ACACTCGGGGCTGTGTCTGGGCTGGTGCCTCTCCATCCTGGCCCCGACTTCTCTGTCAGG 2210

Qy   1240 ------------------------------------------------------------ 1239

Db   2211 AAAGTGGGGATGGACCCCATCTGCATACACGGCTTCTCATGGGTGTGGAACATCTCTGCT 2270

Qy   1240 ------------------------------------------------------------ 1239

Db   2271 TGCGGTTTCAGGAAGGCCTCTGGCTGCTCTAGGAGTCTGATCAGAGTCGTTGCCCCAGTT 2330

Qy   1240 ------------------------------------------------------------ 1239

Db   2331 TGACAGAAGGAAAGGCGGAGCTTATTCAAAGTCTAGAGGGAGTGGAGGAGTTAAGGCTGG 2390

Qy   1240 ------------------------------------------------------------ 1239

Db   2391 ATTTCAGATCTGCCTGGTTCCAGCCGCAGTGTGCCCTCTGCTCCCCCAACGACTTTCCAA 2450

Qy   1240 ------------------------------------------------------------ 1239

Db   2451 ATAATCTCACCAGCGCCTTCCAGCTCAGGCGTCCTAGAAGCGTCTTGAAGCCTATGGCCA 2510

Qy   1240 ------------------------------------------------------------ 1239

Db   2511 GCTGTCTTTGTGTTCCCTCTCACCCGCCTGTCCTCACAGCTGAGACTCCCAGGAAACCTT 2570

Qy   1240 ------------------------------------------------------------ 1239

Db   2571 CAGACTACCTTCCTCTGCCTTCAGCAAGGGGCGTTGCCCACATTCTCTGAGGGTCAGTGG 2630

Qy   1240 ------------------------------------------------------------ 1239

Db   2631 AAGAACCTAGACTCCCATTGCTAGAGGTAGAAAGGGGAAGGGTGCTGGGGAGCAGGGCTG 2690

Qy   1240 ------------------------------------------------------------ 1239

Db   2691 GTCCACAGCAGGTCTCGTGCAGCAGGTACCTGTGGTTCCGCCTTCTCATCTCCCTGAGAC 2750

Qy   1240 ---------------------------------------------------AGGCGTTCG 1248
                                                             |||||||||
Db   2751 TGCTCCGACCCTTCCCTCCCAGGCTCTGTCTGATGGCCCCTCTCCCTCTGCAGGCGTTCG 2810

Qy   1249 GATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCAT 1308
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2811 GATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCAT 2870

Qy   1309 GGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 1368
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2871 GGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 2930

Qy   1369 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 1428
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2931 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 2990

Qy   1429 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 1488
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2991 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 3050

Qy   1489 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 1548
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3051 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 3110

Qy   1549 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 1608
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3111 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 3170

Qy   1609 TAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 1668
          ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db   3171 TAATGGACACGTGGGTGYTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 3230

Qy   1669 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 1728
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3231 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 3290

Qy   1729 GGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTC 1788
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3291 GGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTC 3350

Qy   1789 CCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGC 1848
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3351 CCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGC 3410

Qy   1849 CTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT 1908
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3411 CTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT 3470

Qy   1909 ATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG 1968
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3471 ATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG 3530

Qy   1969 GAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGG 2028
          |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   3531 GAGGGCCTGCCTCACTGGGTCCCAGCTCCCTGCTCCTGTTAGCCCCATGGGGCTGCCGGG 3590

Qy   2029 CTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTG 2088
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3591 CTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTG 3650

Qy   2089 CTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTC 2148
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3651 CTGAGGTGCGTAGCTGCACAGCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTC 3710

Qy   2149 TAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGC 2208
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3711 TAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGC 3770

Qy   2209 CAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG 2268
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3771 CAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG 3830

Qy   2269 TTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAA 2328
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3831 TTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAA 3890

Qy   2329 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCT 2388
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     3591 CTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTG 3650

Qy     2089 CTGAGGTGCGTAGCTGCACAGCTGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTC 2148
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3651 CTGAGGTGCGTAGCTGCACAGCTGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTC 3710

Qy     2149 TAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGC 2208
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3711 TAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGC 3770

Qy     2209 CAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG 2268
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3771 CAGAAGGGCTCCATGCACTGGAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGG 3830

Qy     2269 TTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAA 2328
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3831 TTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAA 3890

Qy     2329 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCT 2388
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3891 CTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCT 3950

Qy     2389 CTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTA 2446
            |||||||||||||||||||||||||||||||| ||||||||||||||||||||  | |||
Db     3951 CTTGCATGGGAGTTTCTAGGATGAAACACTCCACCATGGGATTTGAACATATGAAAGTTA 4010

Qy     2447 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 2506
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4011 TTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGC 4070

Qy     2507 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 2566
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4071 ACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCC 4130

Qy     2567 TTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGT 2626
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4131 TTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGT 4190

Qy     2627 AGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 2686
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4191 AGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGG 4250

Qy     2687 ATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCT 2746
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4251 ATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCT 4310

Qy     2747 TCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCA 2806
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4311 TCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCA 4370

Qy     2807 TCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGG 2866
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4371 TCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGG 4430

Qy     2867 GTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGT 2926
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4431 GTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGT 4490

Qy     2927 TCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 2986
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4491 TCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCA 4550

Qy     2987 AAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCC 3046
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4551 AAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCC 4610

Qy     3047 TGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATC 3106
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4611 TGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATC 4670

Qy     3107 TCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCC 3166
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4671 TCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCC 4730

Qy     3167 TGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATG 3226
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4731 TGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATG 4790

Qy     3227 TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAA 3286
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4791 TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAA 4850

Qy     3287 TGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAA 3330
            ||||||||||||||||||||||||||||||||||||||||||||
Db     4851 TGTTTATGGTGACAAAATTAAAGGCTTTCTTATATGTTTAAAAA 4894
```

```
RESULT 15
AX267728
LOCUS       AX267728                4894 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 702 from Patent WO0173032.
ACCESSION   AX267728
VERSION     AX267728.1  GI:16516400
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
            Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A., Hepler,W.T.
            and Henderson,R.A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0173032-A 702 04-OCT-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..4894
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      928 a   1448 c   1354 g   1163 t      1 others
ORIGIN

  Query Match             62.8%;  Score 2142.8;  DB 6;  Length 4894;
  Best Local Similarity   80.1%;  Pred. No. 0;
  Matches 2872;  Conservative    1;  Mismatches    3;  Indels  708;  Gaps    2;

Qy      455 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 514
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1311 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 1370

Qy      515 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 574
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1371 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 1430

Qy      575 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 634
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1431 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 1490

Qy      635 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 694
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1491 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 1550

Qy      695 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 754
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1551 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 1610

Qy      755 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 814
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1611 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 1670

Qy      815 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 874
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1671 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 1730
```

```
Qy    575 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 634
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1431 CTGAGCCTCTTTCTCATCCCAAGGGCCGGCTGGCTAGCAGGGCTGCTGTGCCCGGATCCC 1490

Qy    635 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 694
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1491 AGGCCCCTGGAGCTGGCACTGCTCATCCTGGGCGTGGGGCTGCTGGACTTCTGTGGCCAG 1550

Qy    695 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 754
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1551 GTGTGCTTCACTCCACTGGAGGCCCTGCTCTCTGACCTCTTCCGGGACCCGGACCACTGT 1610

Qy    755 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 814
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1611 CGCCAGGCCTACTCTGTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTC 1670

Qy    815 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 874
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1671 CTGCCTGCCATTGACTGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAG 1730

Qy    875 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 934
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1731 TGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTG 1790

Qy    935 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 994
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1791 GCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTG 1850

Qy    995 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1054
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1851 TCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTT 1910

Qy   1055 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1114
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1911 CCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCT 1970

Qy   1115 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 1174
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1971 GAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGC 2030

Qy   1175 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 1234
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2031 GAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTAT 2090

Qy   1235 GATGA------------------------------------------------------- 1239
          |||||
Db   2091 GATGAAGGTAAGGCCTTGGCAGCCAGCAGAGGCTGGTGTGGGAGCCGCCCACCAGAGACG 2150

Qy   1240 ------------------------------------------------------------ 1239

Db   2151 ACACTCGGGGCTGTGTCTGGGCTGGTGCCTCTCCATCCTGGCCCCGACTTCTCTGTCAGG 2210

Qy   1240 ------------------------------------------------------------ 1239

Db   2211 AAAGTGGGGATGGACCCCATCTGCATACACGGCTTCTCATGGGTGTGGAACATCTCTGCT 2270

Qy   1240 ------------------------------------------------------------ 1239

Db   2271 TGCGGTTTCAGGAAGGCCTCTGGCTGCTCTAGGAGTCTGATCAGAGTCGTTGCCCCAGTT 2330

Qy   1240 ------------------------------------------------------------ 1239

Db   2331 TGACAGAAGGAAAGGCGGAGCTTATTCAAAGTCTAGAGGGAGTGGAGGAGTTAAGGCTGG 2390

Qy   1240 ------------------------------------------------------------ 1239

Db   2391 ATTTCAGATCTGCCTGGTTCCAGCCGCAGTGTGCCCTCTGCTCCCCCAACGACTTTCCAA 2450

Qy   1240 ------------------------------------------------------------ 1239

Db   2451 ATAATCTCACCAGCGCCTTCCAGCTCAGGCGTCCTAGAAGCGTCTTGAAGCCTATGGCCA 2510

Qy   1240 ------------------------------------------------------------ 1239

Db   2511 GCTGTCTTTGTGTTCCCTCTCACCCGCCTGTCCTCACAGCTGAGACTCCCAGGAAACCTT 2570

Qy   1240 ------------------------------------------------------------ 1239

Db   2571 CAGACTACCTTCCTCTGCCTTCAGCAAGGGGCGTTGCCCACATTCTCTGAGGGTCAGTGG 2630

Qy   1240 ------------------------------------------------------------ 1239

Db   2631 AAGAACCTAGACTCCCATTGCTAGAGGTAGAAAGGGGAAGGGTGCTGGGGAGCAGGGCTG 2690

Qy   1240 ------------------------------------------------------------ 1239

Db   2691 GTCCACAGCAGGTCTCGTGCAGCAGGTACCTGTGGTTCCGCCTTCTCATCTCCCTGAGAC 2750

Qy   1240 ---------------------------------------------------AGGCGTTCG 1248
                                                             |||||||||
Db   2751 TGCTCCGACCCTTCCCTCCCAGGCTCTGTCTGATGGCCCCTCTCCCTCTGCAGGCGTTCG 2810

Qy   1249 GATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCAT 1308
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2811 GATGGGCAGCCTGGGGCTGTTCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCAT 2870

Qy   1309 GGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 1368
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2871 GGACCGGCTGGTGCAGCGATTCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTT 2930

Qy   1369 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 1428
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2931 CCCTGTGGCTGCCGGTGCCACATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGC 2990

Qy   1429 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 1488
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2991 CGCCCTCACCGGGTTCACCTTCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCT 3050

Qy   1489 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 1548
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3051 CTACCACCGGGAGAAGCAGGTGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAG 3110

Qy   1549 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 1608
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3111 CAGTGAGGACAGCCTGATGACCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCC 3170

Qy   1609 TAATGGACACGTGGGTGCTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 1668
          |||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db   3171 TAATGGACACGTGGGTGYTGGAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGG 3230

Qy   1669 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 1728
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3231 GGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGT 3290

Qy   1729 GGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTC 1788
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3291 GGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTC 3350

Qy   1789 CCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGC 1848
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3351 CCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGC 3410

Qy   1849 CTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT 1908
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3411 CTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGT 3470

Qy   1909 ATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG 1968
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3471 ATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTG 3530

Qy   1969 GAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGG 2028
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3531 GAGGGCCTGCCTCACTGGGTCCCAGCTCCCTGCTCCTGTTAGCCCCATGGGGCTGCCGGG 3590

Qy   2029 CTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTG 2088
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGT 1380

Qy       1809 TTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAG 1868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAG 1440

Qy       1869 GCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGG 1928
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGG 1500

Qy       1929 CCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGT 1988
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGT 1560

Qy       1989 CCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTG 2048
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTG 1620

Qy       2049 CTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACA 2108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACA 1680

Qy       2109 GCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGG 2168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGG 1740

Qy       2169 CCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTG 2228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTG 1800

Qy       2229 GAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGT 2288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGT 1860

Qy       2289 TGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCA 2348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCA 1920

Qy       2349 TCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGG 2408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGG 1980

Qy       2409 ATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTG 2466
              |||||||||||||||||||||||||||||||||   | |||||||||||||||||||||
Db       1981 ATGAAACACTCCTCCATGGGATTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTG 2040

Qy       2467 AGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCA 2526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 AGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCA 2100

Qy       2527 CCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGA 2586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGA 2160

Qy       2587 CACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAG 2646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 CACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAG 2220

Qy       2647 CTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCC 2706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCC 2280

Qy       2707 CTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCC 2766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCC 2340

Qy       2767 CCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAA 2826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 CCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAA 2400

Qy       2827 TGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAAC 2886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAAC 2460

Qy       2887 GGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCACTTCCACTCCCC 2946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCACTTCCACTCCCC 2520

Qy       2947 TCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAAC 3006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 TCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAAC 2580

Qy       3007 TTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGA 3066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 TTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGA 2640

Qy       3067 CCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCT 3126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCT 2700

Qy       3127 GTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTC 3186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTC 2760

Qy       3187 TTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGG 3246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGG 2820

Qy       3247 GGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTA 3306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 GGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTA 2880

Qy       3307 AAGGCTTTCTTATATGTTTAAAAA 3330
              ||||||||||||||||||||||||
Db       2881 AAGGCTTTCTTATATGTTTAAAAA 2904

RESULT 14
AX200993
LOCUS       AX200993                4894 bp    DNA     linear   PAT 29-AUG-2001
DEFINITION  Sequence 623 from Patent WO0151633.
ACCESSION   AX200993
VERSION     AX200993.1  GI:15390820
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4894)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Fanger,G.R., Day,C.H., Retter,M.W.,
            Stolk,J.A., Skeiky,Y.A., Wang,A. and Meagher,M.J.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0151633-A 623 19-JUL-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..4894
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      928 a   1448 c   1354 g   1163 t      1 others
ORIGIN

  Query Match             62.8%;  Score 2142.8;  DB 6;  Length 4894;
  Best Local Similarity   80.1%;  Pred. No. 0;
  Matches 2872;  Conservative    1;  Mismatches    3;  Indels  708;  Gaps    2;

Qy        455 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1311 GGCATTGGTCCAGTGCTGGGCCTGGTCTGTGTCCCGCTCCTAGGCTCAGCCAGTGACCAC 1370

Qy        515 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1371 TGGCGTGGACGCTATGGCCGCCGCCGGCCCTTCATCTGGGCACTGTCCTTGGGCATCCTG 1430
```

```
Qy      3007 TTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGA 3066
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2581 TTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACAACCCTGTTTGGAGCTACTGCAGGA 2640

Qy      3067 CCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCT 3126
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2641 CCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCCATCTCAGCCCCCAGAGTATATCT 2700

Qy      3127 GTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTC 3186
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2701 GTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCCCCTGCCTGAGCTAAGGGAGGTC 2760

Qy      3187 TTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGG 3246
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2761 TTATCTCTCAGGGGGGGTTTAAGTGCCGTTTGCAATAATGTCGTCTTATTTATTTAGCGG 2820

Qy      3247 GGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTA 3306
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2821 GGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAATGTTTATGGTGACAAAATTA 2880

Qy      3307 AAGGCTTTCTTATATGTTTAAAAA 3330
             ||||||||||||||||||||||||
Db      2881 AAGGCTTTCTTATATGTTTAAAAA 2904
```

```
RESULT 13
AX267729
LOCUS       AX267729                2904 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 703 from Patent WO0173032.
ACCESSION   AX267729
VERSION     AX267729.1  GI:16516401
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
            Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A., Hepler,W.T.
            and Henderson,R.A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0173032-A 703 04-OCT-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..2904
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      542 a    875 c    773 g    714 t
ORIGIN

  Query Match            64.4%;  Score 2196.4;  DB 6;  Length 2904;
  Best Local Similarity  88.2%;  Pred. No. 0;
  Matches 2560;  Conservative    0;  Mismatches    1;  Indels  343;  Gaps    2;

Qy       770 GTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC 829
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 GTCTATGCCTTCATGATCAGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGAC 60

Qy       830 TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTG 889
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 TGGGACACCAGTGCCCTGGCCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTG 120

Qy       890 CTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCG 949
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 CTCACCCTCATCTTCCTCACCTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCG 180

Qy       950 CTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGT 1009
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 CTGGGCCCCACCGAGCCAGCAGAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGT 240

Qy      1010 CCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAG 1069
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 CCATGCCGGGCCCGCTTGGCTTTCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAG 300

Qy      1070 CTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGG 1129
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 CTGTGCTGCCGCATGCCCCGCACCCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGG 360

Qy      1130 ATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAG 1189
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 ATGGCACTCATGACCTTCACGCTGTTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAG 420

Qy      1190 GGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAA--------- 1240
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 GGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGAAGGCCT 480

Qy      1241 ------------------------------------------------------------ 1240

Db       481 CTGGCTGCTCTAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGACAGAAGGAAAGGCGGA 540

Qy      1241 ------------------------------------------------------------ 1240

Db       541 GCTTATTCAAAGTCTAGAGGGAGTGGAGGAGTTAAGGCTGGATTTCAGATCTGCCTGGTT 600

Qy      1241 ------------------------------------------------------------ 1240

Db       601 CCAGCCGCAGTGTGCCCTCTGCTCCCCCAACGACTTTCCAAATAATCTCACCAGCGCCTT 660

Qy      1241 ------------------------------------------------------------ 1240

Db       661 CCAGCTCAGGCGTCCTAGAAGCGTCTTGAAGCCTATGGCCAGCTGTCTTTGTGTTCCCTC 720

Qy      1241 ------------------------------------------------------------ 1240

Db       721 TCACCCGCCTGTCCTCACAGCTGAGACTCCCAGGAAACCTTCAGACTACCTTCCTCTGCC 780

Qy      1241 --------------------------------GGCGTTCGGATGGCAGCCTGGGGCTGT 1268
                                             ||||||||||||||||||||||||||||
Db       781 TTCAGCAAGGGGCGTTGCCCACATTCTCTGAGGGCGTTCGGATGGCAGCCTGGGGCTGT 840

Qy      1269 TCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGAT 1328
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 TCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGAT 900

Qy      1329 TCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCA 1388
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 TCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCA 960

Qy      1389 CATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCT 1448
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 CATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCT 1020

Qy      1449 TCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGG 1508
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 TCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGG 1080

Qy      1509 TGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGA 1568
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 TGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGA 1140

Qy      1569 CCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTG 1628
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 CCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTG 1200

Qy      1629 GAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCT 1688
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1201 GAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCT 1260

Qy      1689 CCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCT 1748
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 CCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCT 1320

Qy      1749 GCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGT 1808
```

```
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ATGGCACTCATGACCTTCACGCTGTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAG 420

Qy    1190 GGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAA--------- 1240
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GGCGTGCCCAGAGCTGAGCCGGGCACCGAGGCCCGGAGACACTATGATGAAGGAAGGCCT 480

Qy    1241 ------------------------------------------------------------ 1240

Db     481 CTGGCTGCTCTAGGAGTCTGATCAGAGTCGTTGCCCCAGTTTGACAGAAGGAAAGGCGGA 540

Qy    1241 ------------------------------------------------------------ 1240

Db     541 GCTTATTCAAAGTCTAGAGGGAGTGGAGGAGTTAAGGCTGGATTTCAGATCTGCCTGGTT 600

Qy    1241 ------------------------------------------------------------ 1240

Db     601 CCAGCCGCAGTGTGCCCTCTGCTCCCCCAACGACTTTCCAAATAATCTCACCAGCGCCTT 660

Qy    1241 ------------------------------------------------------------ 1240

Db     661 CCAGCTCAGGCGTCCTAGAAGCGTCTTGAAGCCTATGGCCAGCTGTCTTTGTGTTCCCTC 720

Qy    1241 ------------------------------------------------------------ 1240

Db     721 TCACCCGCCTGTCCTCACAGCTGAGACTCCCAGGAAACCTTCAGACTACCTTCCTCTGCC 780

Qy    1241 --------------------------------GGCGTTCGGATGGGCAGCCTGGGGCTGT 1268
                                           ||||||||||||||||||||||||||||
Db     781 TTCAGCAAGGGGCGTTGCCCACATTCTCTGAGGGCGTTCGGATGGGCAGCCTGGGGCTGT 840

Qy    1269 TCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGAT 1328
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 TCCTGCAGTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGAT 900

Qy    1329 TCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCA 1388
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 TCGGCACTCGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCA 960

Qy    1389 CATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCT 1448
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 CATGCCTGTCCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCT 1020

Qy    1449 TCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGG 1508
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 TCTCAGCCCTGCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGG 1080

Qy    1509 TGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGA 1568
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 TGTTCCTGCCCAAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGA 1140

Qy    1569 CCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTG 1628
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 CCAGCTTCCTGCCAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTG 1200

Qy    1629 GAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCT 1688
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 GAGGCAGTGGCCTGCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCT 1260

Qy    1689 CCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCT 1748
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 CCGTACGTGTGGTGGTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCT 1320

Qy    1749 GCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGT 1808
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 GCCTGGACCTCGCCATCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGT 1380

Qy    1809 TTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAG 1868
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TTATGGGCTCCATTGTCCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAG 1440

Qy    1869 GCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGG 1928
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 GCCTGGGTCTGGTCGCCATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGG 1500

Qy    1929 CCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGT 1988
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 CCAAATACTCAGCGTAGAAAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGT 1560

Qy    1989 CCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTG 2048
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 CCCAGCTCCCCGCTCCTGTTAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTG 1620

Qy    2049 CTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACA 2108
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 CTGCCAAAGTAATGTGGCTCTCTGCTGCCACCCTGTGCTGCTGAGGTGCGTAGCTGCACA 1680

Qy    2109 GCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGG 2168
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 GCTGGGGGCTGGGGCGTCCCTCTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGG 1740

Qy    2169 CCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTG 2228
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 CCTTCCAAGGGGGTTTCAGTCTGGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTG 1800

Qy    2229 GAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGT 2288
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801 GAATGCGGGGACTCTGCAGGTGGATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGT 1860

Qy    2289 TGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCA 2348
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861 TGAGACACACCTAGAGAAGGGTTTTTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCA 1920

Qy    2349 TCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGG 2408
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921 TCTCTAAGCCCCTTAACCTGCAGCTTCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGG 1980

Qy    2409 ATGAAACACTCCTCCATGGGATTTGAACATATG--ACTTATTTGTAGGGGAAGAGTCCTG 2466
           |||||||||||||||||||||||||||||||||  | |||||||||||||||||||||||
Db    1981 ATGAAACACTCCTCCATGGGATTTGAACATATGAAAGTTATTTGTAGGGGAAGAGTCCTG 2040

Qy    2467 AGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCA 2526
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2041 AGGGGCAACACACAAGAACCAGGTCCCCTCAGCCCACAGCACTGTCTTTTTGCTGATCCA 2100

Qy    2527 CCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGA 2586
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2101 CCCCCCTCTTACCTTTTATCAGGATGTGGCCTGTTGGTCCTTCTGTTGCCATCACAGAGA 2160

Qy    2587 CACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAG 2646
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2161 CACAGGCATTTAAATATTTAACTTATTTATTTAACAAAGTAGAAGGGAATCCATTGCTAG 2220

Qy    2647 CTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCC 2706
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2221 CTTTTCTGTGTTGGTGTCTAATATTTGGGTAGGGTGGGGGATCCCCAACAATCAGGTCCC 2280

Qy    2707 CTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCC 2766
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2281 CTGAGATAGCTGGTCATTGGGCTGATCATTGCCAGAATCTTCTTCTCCTGGGGTCTGGCC 2340

Qy    2767 CCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAA 2826
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2341 CCCCAAAATGCCTAACCCAGGACCTTGGAAATTCTACTCATCCCAAATGATAATTCCAAA 2400

Qy    2827 TGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAAC 2886
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2401 TGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAGGTAGAGGGTGGGGCTTCAGGTCTCAAC 2460

Qy    2887 GGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCC 2946
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2461 GGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAGCCTGGTTCCCCCCACTTCCACTCCCC 2520

Qy    2947 TCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAAC 3006
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2521 TCTACTCTCTCTAGGACTGGGCTGATGAAGGCACTGCCCAAAATTTCCCCTACCCCCAAC 2580
```